ELECTRIC AND MAGNETIC FIELDS RISK EVALUATION GUIDELINES

California Department of Health Services

December, 1999

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

INTRODUCTION

The California Department of Health Services (DHS) is conducting a program to assess a variety of issues related to electric and magnetic fields. This document explains how one aspect of the multi-part effort will be conducted. It explains how DHS intends to evaluate the potential health risks associated with exposure to electric and magnetic fields. The document has two parts. The first part, guiding principles, explains the background for the effort and the overall approach. It is intended to be accessible to an audience of laypersons who are not technical specialists but who are informed about these issues. The second part, guidance to evaluators, provides guidance for the evaluators who will conduct the review, based on the approach described in general in the first part. It is intended to be accessible to those with technical training or knowledge.

PART ONE: GUIDING PRINCIPLES

This first part of the guidelines for the evaluation of potential health risks associated with exposure to electric and magnetic fields (EMFs) provides background about the origins and purposes of the evaluation, explains how it fits into a larger project related to EMFs, summarizes the overall approach of the evaluation and presents some of the most important substantive elements of the process.

I. The California EMF Project and the Process for Developing Risk Evaluation Guidelines

The State of California is in the midst of a large project that is examining from a variety of perspectives the significance to human health of exposure to electric and magnetic fields. The California Department of Health Services (DHS) is the lead agency coordinating this review. DHS is also conducting additional research on health effects that may be associated with exposure to EMFs. The California Public Utilities Commission (CPUC) has directed investor-owned utilities to provide funding for the work. Municipal utilities are providing funding voluntarily.

The CPUC has jurisdiction over all investor-owned electric distribution lines in the state. The new Independent Systems Operator has jurisdiction over most of the transmission lines formerly controlled by investor-owned utilities. After a consensus process and a series of hearings in the early 1990s the CPUC announced in November, 1993 a policy of low-cost and no-cost mitigation to reduce exposures to electric and magnetic fields for new construction and of taking no action on existing facilities. The CPUC has sponsored additional research and evaluation, including this project. The project began in 1994 and is to be completed in the year 2001.

The California EMF project has several elements. These include the following:

- School Exposure Assessment – analysis of EMF exposure in schools, ways that EMFs could be mitigated and cost of such mitigation alternatives.
- School, Power Grid and Land Use Policy Analyses – evaluation of policies related to schools and low-voltage distribution lines and high-voltage transmission lines. Policies being evaluated are options for retrofitting schools and the power grid and land use policies for areas adjacent to distribution and transmission lines. Decision trees that describe costs and benefits of various policy options in quantitative terms are being prepared. The reports will also explicitly discuss ethical considerations.
- Public Health Risk Evaluation – review of the evidence for health effects associated with EMF exposure and evaluation of the likelihood of effects for people in California. This document pre-

sents guidelines for how to do the public health evaluation. The final evaluation by staff of DHS and their consultants will be reviewed by the Science Advisory Panel and made available for public comment.

- Worker Exposure Study – development of a method for evaluating likely occupational exposure to EMFs by looking at activities and equipment used in different jobs.
- Electric Car Study – in cooperation with the Federal Department of Transportation, which carried out measurement of EMFs in electric vehicles.
- Miscarriage Study – one thousand women in San Francisco are included. EMFs to which they are exposed are being measured. The women will be followed through pregnancy to see if exposure to EMFs is related to miscarriage.
- Policy Integration – the final product will integrate the other elements into broad policy options and when appropriate into recommendations. These can be used by the state and local boards of education, the CPUC, local governments, state and local health agencies, private individuals, and others to decide what if any action to take to reduce or prevent exposure to EMFs.

Role of These Guidelines in the Overall Project

One of the reasons that this project has been undertaken is because people are concerned about the possibility that EMFs may cause health effects. Whether they cause health effects has been very controversial. At present, there is no consensus on this question. DHS plans to independently evaluate the potential health risks associated with exposure to EMFs. This evaluation will be used in the policy analyses developed for schools, power grids, and land use. The health risk evaluation will provide information that will be used to define the benefits, if any, of policy options that reduce exposures to EMFs. It will also be used in a final policy integration document.

This document proposes guidelines for how DHS would conduct the risk evaluation. DHS is presenting these guidelines for review and discussion in advance of the evaluation itself. The guidelines and the evaluation itself are being developed by scientists and will be subjected to peer review. In addition, a wide variety of perspectives of stakeholders and other representatives of the public is being sought to make the evaluation process as useful as possible for a wide range of purposes.

To achieve this, the guidelines describe how the risk evaluation will interpret evidence to determine how likely it is that EMF exposure causes adverse effects. We have tried to define our terms and to use plain language. DHS hopes that securing a broad understanding of the risk evaluation guidelines beforehand, with careful attention to the logic of risk evaluation and to the likely application of the evaluation, will make the results more useful.

Process to Develop the Guidelines

The California Department of Health Services has taken several steps to gain input before preparing this draft. Since the beginning of the EMF project, DHS has been working with Stakeholder Advisory Consultants (SAC), who include representatives of interests affected by EMFs. The SAC has provided advice throughout the project, reviewed requests for proposals, and helped to design the process for development and review of these guidelines.

To gain advice from experts in the field of risk evaluation, in 1997, DHS commissioned a team of consultants to develop recommendations on how to conduct the risk evaluation. The expert team, known as the Worcester Group, submitted its report in October 1998. The report included a wide-ranging discussion of issues associated with the evaluation as well as specific advice on how to deal with some of these issues. DHS also hired a consultant to assist in shaping the risk evaluation guidelines. DHS and its consultant drew upon the Worcester Group report, along with previous comments and perspectives from members of the SAC, in developing these guidelines.

DHS established a Science Advisory Panel (SAP) comprised of experts in several relevant disciplines, including toxicology, epidemiology, ethics, physics, and statistics, to review the risk evaluation guidelines (Appendix 1). DHS selected the members of the SAP after review of possible candidates by the SAC. The SAP met to discuss a draft of these guidelines.

This draft incorporates comments and suggestions from members of the SAP and the SAC. While neither the SAP nor the SAC are unanimous in their views on many issues, this draft reflects DHS' best attempt to integrate as many of the varying perspectives as possible.

The risk evaluation guidelines have been reviewed twice. The SAP reviewed an agency draft and discussed it at a meeting on February 22, 1999, in Oakland. The draft was extensively revised, through consideration of comments from both the SAP and SAC and was then sent for full review within DHS. DHS released the guidelines for public review on July 15, 1999. Comments were due by August 31. Comments were received from 28 individuals, including several members of the SAC and SAP. These have been thoroughly reviewed. This draft reflects many changes based on these comments. A summary and response to comments has also been prepared as a separate document.

This final version of the guidelines will be either accepted or rejected by the SAP. That will be the final step in the preparation of the guidelines.

Conducting the Risk Evaluation for EMFs

Once the guidelines are adopted, DHS staff and its consultants will conduct the risk evaluation for EMFs. The results of the evaluation will then be included with other elements of the program in the policy integration step.

The steps in the development of the risk evaluation based on these guidelines are below. Dates are not yet associated with these steps.

Initiate risk evaluation process:

- develop and review pro and con arguments and supporting statements
- conduct internal workshop on the risk evaluation to clarify weight of evidence and derived degree of confidence and assign International Agency for Cancer Research categories
- prepare first draft of the risk evaluation (Program staff)
- Science Advisory Panel (SAP) review of the first draft of the risk evaluation
- prepare second draft of the risk evaluation, incorporating the comments of the SAP
- agency review
- public and SAC review of the risk evaluation
- prepare final draft of the risk evaluation, incorporating public comments
- SAP review of the final draft of the risk evaluation.
- SAP final meeting for consideration of the risk evaluation (projected for 2001)

DHS recognizes that it would be best to update the evaluation periodically in the future as significant new findings emerge from scientific study. At present, funding is not available for such an effort, but a proposal for periodic review may be included in the risk evaluation.

Purpose of and Audience for this Document

This document has been prepared for two audiences. Part One, "guiding principles," explains the rationale for the approach proposed for assessing risks of EMFs. The intended audience is stakeholders who wish to provide input on how the evaluation will be used and how its information should be "packaged" for use by decision-makers. Part Two provides "specific guidance" to those who will conduct the evaluation. These will be employees of DHS and their consultants. Part Two uses more technically oriented language.

II. Uses of a Public Health Risk Evaluation

Stakeholders make decisions in a variety of contexts about EMFs. People with different responsibilities in different organizations make decisions. Stakeholders have different uses for an evaluation of potential health concerns. This evaluation is intended to respond to as many of these varying contexts and purposes as possible.

The focus is on evaluating health risks. Some of the concerns expressed by stakeholders about other aspects of EMFs (such as loss of property values) will not be addressed in the public health evaluation, but will be addressed in other parts of the EMF project and in the overall policy integration.

DHS recognizes that the views of people interested in these issues may have solidified in some cases. Other entities have conducted reviews of evidence of whether EMFs cause health effects.[1,2] This evaluation is being designed to take a fresh look at the evidence using a process that is defined in advance with the advice and participation of stakeholders. This project will address decision-makers in California. It may be of use elsewhere as well. While the available evidence is likely not to be sufficient to resolve all uncertainties about any health effects associated with EMF exposure, it is important to come to closure on interim policy based on what we know now.

DHS has identified four ways in which the evaluation is likely to provide useful information:

- Identification and characterization of potential health risks, if any, in new and existing schools and ways to address them. This could contribute to policy recommendations for the Department of Education and local school boards.
- Identification and characterization of potential health risks, if any, from new and existing home grounding systems, power transmission and distribution lines and ways to mitigate exposure. This could contribute to policy formation by the CPUC, elective boards that oversee municipal utilities, electric utilities and the Legislature.
- Identification and characterization of potential health risks, if any, from products, electric vehicles, and appliances. This could contribute to recommendations to the public about personal exposure to EMFs. Individuals and public and private organizations may make use of this information in their own decisions.
- Identification of health risks and ways to address them for consideration by state and local bhealth departments.

Differing Contexts Have Differing Needs for Confidence

DHS recognizes that a fundamental challenge for this evaluation is that scientific evidence may not allow for certainty in conclusions about health risks. Specifically, DHS recognizes that scientists may or may not be able to conclude that it is more than 50% likely that exposure to EMFs causes various diseases. Nonetheless, we will do our best to characterize our degree of confidence and our uncertainty about it. To facilitate the policy analysis we will also characterize the theoretical size (magnitude) of any risks if they were real.

This approach is appropriate because decisions in different contexts have different needs for certainty. In some contexts, a high degree of confidence that exposure to a potentially harmful agent causes adverse effects is needed before action is taken. In other contexts, less confidence is needed.

Types of decisions that are usually based on a high degree of confidence include:

- Actions by government agencies to reduce or prevent exposures to agents that may pose risk. Public agencies usually require a high degree of confidence that something is a hazard before requiring reductions in exposure.
- Mandatory warnings to the public.
- Remedies imposed through litigation. Civil courts often use a "more likely than not" standard for proof that harm resulted from an exposure. Criminal courts require the more stringent "beyond a reasonable doubt" standard for criminal sentences.

Types of decisions that may be based on a lower degree of confidence include:

- Some mandatory warnings on pharmaceutical products about potential risks. For example, warnings that pregnant women may experience harm are often required to be included with drugs even if the certainty that this will occur is low.
- Voluntary actions to avoid exposure. Individuals may choose to avoid exposure even if their certainty of harm is low, especially when the cost of avoiding the exposure is also low. People may decide to avoid use of devices that create high EMF exposures or to ask their contractors to use wiring practices that produce relatively low EMFs, for example.
- Voluntary warning of customers about risks. The decision to voluntarily warn or protect customers may occur with lower degrees of certainty when ethical concerns are salient, costs are low, or risks of litigation are high.
- Funding research about risk. Funding agencies often award research monies to study a potential source of risk before the risk is proven.

The Public Health Risk Evaluation Aims to Accommodate Many Styles of Risk Management

This evaluation will first use an approach similar to that used in risk assessments of environmental agents which are prepared for regulatory agencies to describe the likelihood that those agents cause health effects. In addition, we will also use an approach that is more explicit about our degree of confidence that exposure to EMFs causes disease.

In regulatory contexts, risk assessors do not typically quantify their degree of confidence that an agent poses a hazard, but rather use a weight-of-evidence approach to classify agents into categories. For example, the US Environmental Protection Agency classifies compounds as "known" carcinogens (class A), probable carcinogens (class B), possible carcinogens (class C), as having insufficient evidence to classify as a hazard (class D) or as having no evidence of carcinogenicity (class E). They do not provide a quantitative estimate of their degree of certainty.

Regulatory agencies seldom take action to reduce exposure to agents if carcinogenicity is considered only "possible" or if little is known. Regulators may defer actions to reduce or prevent risk until more information accumulates. Generally, regulatory action is taken for carcinogens classified as probable or known human carcinogens, though some actions, including development of drinking water advisories, have been taken for chemicals considered "possible" carcinogens.

Alternatively, risk managers can react to limited knowledge by proposing no- and low-cost actions to reduce risks. For example, the CPUC (and also the Swedish government) have recommended a "no- and low-cost avoidance" approach to new powerline construction. This means that they would build new power lines in a way that would reduce exposure, but that would not increase costs significantly or at all. The California Department of Education requires buffers between new schools and power lines. Another example comes from the policy debate over release of gases that may contribute to global warming. Many policy analysts have suggested that increasing energy efficiency would reduce release of these gases while also decreasing costs and should be adopted even if it is uncertain that climate change is occurring. Such policies are often called "no regrets" policies. DHS plans to consider such policy options in its overall EMF project and to do this using the tool of decision analysis.

Evaluating courses of action with decision analysis requires the risk evaluator to quantify the degree of confidence that a hazard exists and to estimate the magnitude of the hazard, if real. This makes it possible to evaluate a range of options and to determine if there are courses of action that might otherwise not have been identified. It may also show that popular solutions are not advisable.

The disadvantages of the approach are twofold. First, decision analysis is highly technical and not readily understandable by anyone without specialized training in quantitative research methods. Second, the estimated degree of certainty and magnitude of potential risk numbers used could take on an aura of reality that comes to dominate public perception. Framing action using hypothetical numbers may be perceived very differently by many members of the public than explaining any action as being based on limited knowledge. Action based on "limited knowledge" may be perceived this way: "We weren't sure there was any hazard at all, but just to be careful we took precautionary actions." The action based on a hypothetical number may be perceived in another way: "This hazard was killing x people a year, so we had to take precautionary action."

Because we concluded that decision analysis could be informative, the California EMF program has funded quantitative decision analysis. Our risk evaluation will provide and justify numbers for this analysis. But we are committed to presenting our evaluation in ways that allow individuals, private sector decision-makers, the CPUC, and local boards of supervisors to use any style of risk management and risk communication they choose. Our mode of risk evaluation will strive to accommodate all these risk managers.

The public health risk evaluation will have a number of products intended to be useful to different decision-makers:

- A hazard identification using customary categories for weight of evidence for carcinogens developed by the International Agency for Research on Cancer (IARC), as proposed in the World Health Organization risk assessment for EMFs. This approach would be applied to cancer and non-cancer outcomes (see Appendix 3).

- A description of our degree of confidence that EMFs cause various diseases using language presented below.
- For decision-makers who make judgments based on the coherence of evidence we will present pro, con and summary arguments for whether EMFs cause the diseases evaluated.
- For decision-makers who want to make decisions about further research on EMF, if any, or to delay action while waiting for more information, we will describe the state of the science and whether there are important studies in the pipeline. We will provide pro and con arguments and summary opinions on whether certain lines of investigation are likely to provide positive or negative breakthroughs and how long research funding would be needed before results were forthcoming.
- For those decision-makers concerned about the potential for unequal vulnerability of sub-populations or unequal distribution of exposure we will review the evidence for both of these as it relates to EMFs.
- We will also provide a "recommended risk communication statement" acknowledging different ways the degree of confidence about the risks of EMFs can be legitimately framed.
- For the quantitative decision analysis we will provide a degree of confidence that EMF exposure causes diseases and an estimate of the magnitude of risk, if real. Specifically, we will answer these questions:

  What is our degree of confidence that the range of usual environmental and/or occupational exposures to EMFs is a contributing cause that partially explains the epidemiological associations seen with certain diseases? (Answer: We are virtually certain that smoking two packs a day of cigarettes causes lung cancer. We are virtually certain that drinking two liters of water a day causes no adverse effects.)

  If EMFs caused one or more of these diseases what is the magnitude of the added lifetime risk conveyed by the range of EMF exposures? (Answer: About 10% of people who smoke two packs of cigarettes a day eventually get lung cancer. About one in a thousand non-smokers who live for a long time with a smoker will develop lung cancer they would not have gotten otherwise.)

  How much can we reduce the probability of harm through mitigation that reduces exposure to the attributes of EMFs? (Answer: Stopping smoking cigarettes cuts the lifetime lung cancer risk of heavy smokers from about one in ten to close to zero and of "secondhand smokers" from one in a thousand to close to zero, but removing the nicotine would not affect cancer risk per se.)

  How many cases of disease could be prevented each year in California by reducing current exposures to the suspected bioactive attributes of EMFs? For each disease, we will include a statement of the best estimate of the current incidence of the disease, the number of cases that might be expected to result from the exposures experienced by the people of California, and an estimate of the increase that this represents over the baseline. (Answer: A hundred thousand lung cancer deaths each year and about a thousand from secondhand smoke could be avoided in the US by eliminating cigarette smoking.)

Terminology to Describe Degree of Confidence

As noted, evaluators will be asked to frame in two ways their conclusions about whether EMF exposure causes disease. First, they will apply classification systems developed by the International Agency for Research on Cancer (the IARC categories are shown in Table 2 in Part Two). They are the same as used by the National Institute of Environmental Health Sciences (NIEHS) in their 1998 risk assessment and will be used by the World Health Organization for their future EMF risk evaluation (see also Appendix 3). Second, evaluators will be asked to give his/her degree of confidence as to whether associations between EMF exposure and disease are causal in nature.

To assist in defining this degree of confidence, the DHS scientists responsible for this evaluation will receive training in "probability elicitation." For each disease, each member of the evaluation team, after a structured and thorough discussion, will express his or her degree of confidence that the epidemiological associations seen are causal in nature. After this they will consider the size of the effect if real. This two-step elicitation reflects the structure used in the two policy projects in the EMF project. On the basis of the discussions the evaluators will select an appropriate narrative description using the terminology in Table 1. This table provides suggested ways of describing degrees of certainty for relationships considered during the evaluation. The evaluation team may decide that fewer categories are appropriate in some or all cases.

Table 1. Proposed language for describing degree of confidence in EMF causation of disease

narrative description	percent confidence
virtually certain to be a cause of a particular disease	>98
highly probable that it is a cause	≥90
possibly a cause—more than 50% likely	>50 and <90
possibly a cause—less likely than 51%, but not very improbable	≥10 and ≤50
(very) improbable that it is a cause	>2 and <10
virtually certain not to be a cause	≤2

To deal with the reality that lack of evidence, poor technical quality of evidence, or conflicting evidence can make it difficult to specify one's degree of confidence, the evaluators may comment on the quality of the evidentiary base and will give a range for their degree of confidence.

How Big is the Effect if the Epidemiological Associations Are Real?

It is one thing to say one is convinced that an agent causes some cancer at doses found in the everyday environment. It is more difficult to go to the next step and specify the added risk conveyed by a particular environmental dose, or to estimate the number of cases of disease which are attributable to the range of environmental exposures now found in the population.

Compared to some environmental agents, we have a large amount of information about the population's range of exposure to at least one aspect of the EMF mixture, the 60 Hz field average over time (the “time-weighted average”or TWA), at home, at work and elsewhere. This information comes from special surveys and epidemiological studies that have used computerized personal monitors which took readings every few seconds. We also have a good idea as to the proportion of the population who work in various job categories and those whose residences fall in different “wire codes” (a way of classifying powerlines as to current flow and proximity to houses). These measurements have been associated with disease in some epidemiological studies.

We can calculate the added risk, if real, from being above exposure levels used in epidemiological studies or from living in a house with a particular wire code or in a particular job classification. One can also calculate the theoretical impact on the overall disease rate if everyone occupied the exposure level or the wire code or job category with the lowest apparent risk

It is more difficult to estimate the impact of changing exposures at levels other than those studied by epidemiologists. Estimating any dose-response relationship for EMFs is also difficult. We are proposing to examine this issue in the evaluation, though we recognize that data may be available for only a few diseases.

Our power grid policy analysis has been designed by consultants to this project and has the capacity to evaluate mitigation using certain assumptions about the dose-response relationship. The models require certain specific inputs from the risk evaluation, and the evaluation will be conducted to supply these. However, evaluators will also be free to examine all models that they feel are appropriate and to come to whatever conclusions they believe are justified about whether available data supports a model.

The risk evaluation will discuss whether there is anything in the various kinds of evidence that would allow favoring TWA or one type of dose response over another. Of the various diseases that we propose to study, some may have enough exposure information to begin addressing these issues. Others may not. The risk evaluation must discuss whether dose response evidence for one disease is valid for another disease.

As described in Part Two, we will present a range for:

- the theoretical accumulated risk from a lifetime at the 90th percentile of exposure
- the attributable population burden derived from the current distribution of exposure in the population

EMFs as a Mixture of Attributes

EMFs have many attributes, including frequency, intensity and polarization. EMFs from different types of sources may have different combinations of attributes. Remediation options may change some of the attributes but not others. We do not know yet which if any of these several attributes singly or in combination are important in causing health effects.

Environmental levels have been measured as time-weighted average (TWA) values for typical time periods at home or at work. They exhibit a strongly skewed distribution, with median values around 1 milliGauss (mG) in the residential environment and 1.5-2 mG in most occupational environments, but are sometimes measured at several milliGauss in residential environments and tens to hundreds of milliGauss in the most exposed occupations.

To be helpful a risk evaluation should discuss (a) whether study aspects are well correlated with the 60 Hz TWA magnetic field strength, which has been associated with disease in some epidemiological studies. It should also discuss (b) the strength of the evidence that links various aspects to biological effects or disease.

Uncertainty about which attribute of EMF may be associated with adverse health outcomes has been advanced as a reason to delay remedial action regardless of whether the EMF mixture is determined to be hazardous. A mitigation action, it is said, might modify the wrong attribute or lower one inactive attribute of the mixture and increase a harmful attribute. In a special appendix separate from the risk evaluation we plan to discuss the impact of various proposed mitigation options on the various attributes of the EMF "mixture" and assess how their efficacy could be affected by this uncertainty. For example, what if the TWA were only correlated with some other aspect of the magnetic field that did not always respond to mitigation that targeted the TWA? What if it were correlated to the square of the rms field, as argued by Adair[3] and Wilson[4]?

Terminology for Patterns of Evidence

In describing a body of evidence we want to avoid using adjectives that presuppose policy directions. We plan to use the following terminology.

- To describe relationships between exposures to EMFs and all types of outcomes we will use the terms: "increase in occurrence," "no change in occurrence," or "decrease in occurrence." The term "occurrence" can refer to any measured outcome.

We will include in the review individual studies that reported results which didn't reach conventional statistical significance, since a barely detectable association based on the size and quality of the study may only become apparent in a meta-analysis (statistical technique for combining results from many studies) or a less formal equivalent review. We will provide confidence limits for individual studies or calculated "probability" values when these are available. There is controversy about depending upon statistical tests to evaluate or screen studies. We will look at the evidence both ways and comment on whether this alters the conclusions. (Where we describe tests of significance we will prefer two-tailed 95% confidence limits or when only p values are available we will specify if they are one or two tails, with preference for two-tailed tests.)

- To describe outcomes that are observed always or almost always in repeated experiments or studies, we will use the word "consistent."

- We will characterize as "recurrent" those outcomes that while not always seen are observed repeatedly in studies and have no clear alternative explanation.

It is not uncommon for agencies in their summary statements after a risk assessment to characterize the strength of an association, not as a number with confidence limits, but as "strong" or "weak." We will use instead terms which are policy-neutral. The terms "strong" and "weak" have several quite different interpretations, so in public summary statements we will use phrases like these, which express more clearly what we have in mind:

- To express whether a finding is worthy or unworthy of societal or policy concern, "The magnitude of theoretical attributable lifetime risk (for cancer) is larger/smaller than the one per 100,000 level that triggers notice under Proposition 65."

- To express whether a finding is easily or barely detectable given the size and quality of the scientific studies used, "The difference of occurrence between exposed and unexposed individuals was easily, barely, or not reliably detectable given the size and quality of the studies available."

- To express whether an association is large or small compared to some other association, "The added risk or proportion of total cases of disease x attributable to EMFs is larger, same or smaller than the added risk or proportion of total cases of disease x attributable to agent y."

It should be noted that even barely detectable effects from many epidemiological studies can be larger than those that would call for notice under Proposition 65 in California.

Since "robust" can also have multiple interpretations we will avoid its use and instead say:

> "The size of the effect was easily detectable given the size and quality of the study, was seen consistently in repeated experiments and was larger than the variation between the various control groups."

We wish to avoid the ambiguity of such statements as "there is no evidence that x causes y," which could mean that there are no studies on this topic or that there are plenty of studies but all of them fail to show that x causes y. We will therefore talk about the "evidentiary base" to describe the volume of evidence and will characterize it as "absent," "scant," "moderate" in size or "voluminous." We will talk about the "pattern of evidence" to denote the results in that evidentiary base. So we might say, "There is no evidentiary base to address the question of whether x causes y," or, "There is a voluminous evidentiary base on whether x causes y, and the pattern of evidence consistently suggests that x does not cause y."

Dealing with Study Quality and Describing It

We intend to review studies that have been published or accepted for publication. For studies the California EMF program has sponsored, we will include those that have passed the external peer review which we have arranged, even if the study has not yet been submitted for publication.

Epidemiologists tend to think about quality issues differently from experimentalists. Since epidemiologists rarely perform experiments (randomized trials are the exceptions) they rarely can eliminate bias and confounding and measurement error to the degree which is possible in an experiment. The experimentalist tries to control everything and will often discard a study entirely if there was a failure to control any of the desired parameters. The experimentalist tends therefore to think in terms of "good quality studies" and "bad quality studies" and simply ignores the latter category. The epidemiologist does not have this luxury and tends to evaluate the direction of the biases introduced by the inevitable lack of perfection in study designs. Although we will acknowledge standard experimental practice and whether an experimental study was carried out under standard, "good laboratory practices" when discussing experimental studies, we will also discuss the expected direction of bias, measurement error and confounding in both experimental and epidemiological studies. The structured questions in Section Two assure that these issues are explicitly dealt with.

Avoiding Conflict of Interest

The DHS scientists involved in the assessment and their consultants will be asked to complete the standard California conflict of interest disclosures. Scientists with conflicts of interest will be excluded from the review team. The members of the Scientific Advisory Panel are free of financial conflict of interest and have not been involved in the EMF controversy.

Explaining "Degree of Confidence" and "Magnitude of Risk" to the Public

This way of talking about the evidence can be illustrated by applying it to the evidence related to the carcinogenicity of benzene, arsenic and ferric oxide.

> Benzene: The US EPA and CalEPA have classified benzene as a known human carcinogen on the basis of a voluminous evidentiary base of acceptable quality in animals and a number of occupational studies of acceptable quality in humans that show an easily detectable increase of cancer occurrence given the strength and weaknesses of the studies. Scientists at DHS think it is somewhere between

more than 50% certain, but less than virtually certain that benzene in typical urban air could increase the rate of leukemia in the population to some degree. However, the upper bound of theoretical increase in occurrence would be well below the power of the best epidemiological studies of the general population to detect. The upper bound of theoretical risks from a lifetime of exposure would be on the order of 10 per 100,000 and is of regulatory concern since California regulates at the level of one in 100,000 theoretical lifetime risk. The chance of escaping leukemia after a lifetime of breathing benzene in urban air would be 99,990 per 100,000, so the individual risk is small. Some people want to know what proportion of the total burden of disease in the population is attributable to a factor like benzene in urban air. The total lifetime risk of leukemia from all causes is about 700 per 100,000. Thus, benzene in air would not account for much of the total leukemia rate in the population.

Arsenic: The US EPA and Cal EPA have classified arsenic as a human carcinogen based on a voluminous evidentiary base of human occupational and drinking water epidemiology which includes good quality studies showing effects easily detectable given the size and quality of the studies and despite an adequate evidentiary base in animals which until recently failed to experimentally demonstrate cancer in animals. DHS scientists believe that it is highly probable to virtually certain that arsenic in occupational settings and in drinking water can produce some cancer. Epidemiological evidence suggests that in some parts of California with high arsenic content in water the lifetime theoretical risk could reach 1,000 per 100,000, far above the one per 100,000 regulatory level. Even in these areas an individual would have a 99% chance of escaping cancer caused by arsenic. We do not have sufficient exposure information about the general public to estimate the excess of cancer caused by arsenic.

Ferric oxide: Based on an adequately voluminous evidentiary base in animal studies which have not shown an increased occurrence of tumors in animals and a number of occupational studies in humans which have not shown an increased cancer rate when other known carcinogens were absent from the work place, the International Agency for Research on Cancer has said this agent is "not classifiable as to human carcinogenicity and with animal evidence suggesting lack of animal carcinogenicity." DHS scientists would estimate that ferric oxide is very unlikely to virtually certain not to cause cancer in occupational or environmental settings.

IV. Evaluating Streams of Evidence

There are four principal types of evidence that are relevant to this review—biophysical theory, animal and human studies of biochemical and physiological changes (mechanistic studies), animal studies that focus on disease, and epidemiology. A fundamental challenge for this evaluation is to review and make sense of these four different types of evidence. The guidelines explain how these different types of information will be considered. They explain the questions that evaluators should consider for each type to ensure that all relevant issues are considered..

As a general rule, a pattern of positive and negative results in a body of evidence will incriminate an agent as hazardous if that kind of pattern was more likely if the agent were indeed hazardous than if the agent was not hazardous. That is to say, one is influenced by the relative likelihood of the pattern of evidence and the quality of the evidence that is displaying this pattern. The quality of evidence is also important.

It may be helpful to describe the pattern of evidence that would make us virtually certain that EMFs cause disease and the pattern of evidence that would make us virtually certain that they do not. Completely convincing evidence would include associations between exposure and disease in epidemiology easily detectable by the available studies. Epidemiological studies for all alternative explanations would show no change in occurrence, tests for bias would show no bias. Diseases would be strongly induced in two species of experimental animals at environmental levels of EMFs. The mechanism linking exposure to the first molecular event would be clearly identified in several experiments, and biophysical theory would predict the observed response to exposure. We would have identified the attributes of the EMF mixture that cause these effects.

Even if EMFs were hazardous, the likelihood of such a clear evidentiary pattern would be extremely low. Few if any recognized hazards boast such a clear pattern, but we present this extreme case to make our point: the relative likelihood would be a big number because the likelihood of this pattern of evidence by chance alone is vanishingly small.

We can also describe evidence that would be completely convincing that there is no effect. Sufficiently large and well-designed epidemiological studies would not detect effects. Further study would show that biases or confounders explain previously reported associations between exposures and disease. Studies in animals using a number of plausible attributes in the mixture would not detect effects even in large experiments at exposures higher than those typically found in the environment (but lower than those known to cause acute effects). The positive results in experiments to date would be shown to be due to factors such as temperature or vibration. The physical induction mechanism of more intense EMF effects would be understood. Theory would explain the threshold, far above everyday exposures, below which effects would not occur. Experiments would confirm these predictions.

Of course, most "safe" agents don't boast a pattern of evidence which is as clear and comprehensive as the one described above, but we present this extreme case to make our point: the relative likelihood of this pattern of evidence would be a very small, fractional number since the likelihood of this pattern of evidence occurring if EMFs caused disease would be much smaller than the likelihood of this pattern if they didn't.

When research results do not converge to a pattern of evidence which clearly builds confidence or clearly reduces confidence that there is a hazard and there have been a number of research iterations exhausting all reasonable avenues of investigation, one has reached the point of research exhaustion, the point where evidence has been shown to be unhelpful. It is important to determine if one has reached that point with EMFs.

The Challenge in Combining Evidence

The relevant evidence includes studies of variable strength and quality that must be considered together to reach a conclusion. Answering the questions below would summarize the overall pattern of combined evidence.

Biophysical theory: Does theory predict that the usual range of residential EMFs would affect normal biological processes? If not, does theory predict that occupational levels higher than the residential average, but lower than those at which effects are clearly explained by the well-understood mechanism of induced currents would affect normal biological processes?

Mechanistic research: Have normal biological processes been affected by residential levels of EMFs? If not, have normal biological processes been affected by higher levels of EMFs? If biological processes have been altered, do these steps lie on a causal chain to disease? Are these diseases related to those seen in epidemiological studies? If changes occur, are they likely to be reversed or repaired upon cessation of exposure?

Whole-animal studies focusing on disease: Have residential or occupational levels of EMFs caused disease in animal experiments? If not, have EMFs at levels higher than residential or occupational averages but lower than those at which effects are clearly explained by the well-understood mechanism of induced currents caused disease in animal experiments? Do these findings demonstrate a mechanism for effects of EMFs at higher levels of biological organization? Do these findings demonstrate effects that are relevant to humans? Are the animal effects what would be expected from mechanistic studies of EMFs?

Epidemiology: Do epidemiological studies show an increase in occurrence of disease, a decrease in the occurrence of disease, or no change in the occurrence of disease to be associated with exposure to EMFs? Is the magnitude (easily, barely) detectable by the size and quality of studies performed? If so, are changes due to confounding or bias? Are the effects consistent or recurrent?

Answering yes or no to all these questions could generate a logic tree with hundreds of branches. The patterns of evidence mentioned above that would build our degree of confidence close to 100% or would decrease it towards zero would be the most extreme, outermost branches of this logic tree. Of course, it would be the rare agent which would have a pattern of evidence as extreme as either of these outer branches. Deriving a degree of confidence from the patterns of evidence represented by the many inner branches of the logic tree presents a bigger challenge. This requires considering the likelihood of the observed pattern if EMFs were hazardous relative to the likelihood of the observed pattern if EMFs were not hazardous. Moreover, one would need to consider the quality of the evidence displaying the observed pattern.

Using Evidence to Estimate Degree of Confidence

It is possible to turn to probability theory for approaches to the problem of combining evidence and describing one's degree of confidence in associations between exposure and effect. This type of approach is often referred to as a "Bayesian" approach to scientific reasoning. [5] This method uses the concepts of probability to compare one's initial amount of confidence in a hypothesis to the confidence one has after considering more evidence. How likely would the pattern formed by the new evidence be if the hypothesis were true? how likely if it were not true? Comparing the strength of the likelihood in each case tells how influential the new evidence is, thus modifying the "degree of confidence." The Bayesian view allows for evidence that can strengthen or weaken our degree of confidence. We believe that it is a reasonable way to conceptualize scientific practice.

As is explained more completely in Appendix 2, one can conceive of types of evidence as falling into four classes, described below. The class is determined by statistical power, degree of measurement error, and control of confounding and bias.

Uninformative: This type of evidence is so weak that no matter what result you get from it the likelihood of that result if EMFs were hazardous is about the same as it would be if EMFs were not hazardous, so no result will change your degree of confidence much. Using the language of laboratory tests, this kind of evidence is neither "sensitive" (high "true positive rate") nor "specific" (low "false positive rate") (for people unfamiliar with these terms see glossary and Appendix 2).

Strengthening and weakening: If EMFs were indeed hazardous this type of evidence is very likely to give a positive result and would be much more likely to give a positive result than if EMFs were not hazardous. Therefore, a positive result would really strengthen your degree of confidence. If EMFs were hazardous, this type of evidence is quite unlikely to give a negative result and is much less likely to give a negative result than if EMFs were safe, so a negative result would really weaken your degree of confidence. (An example of strengthening and weakening evidence from another domain might be studies attempting to link lung cancer to cigarette smoking. Our ability to measure the intensity and duration of smoking is pretty good, our ability to control confounding factors is good and since the effect is large compared to the statistical power of economically feasible studies, the overall ability to detect the effect is good. Therefore, the likelihood of a positive result is quite large if cigarettes are hazardous and is quite small if cigarettes were safe. Therefore, a positive result strengthens one's degree of confidence about the dangers of cigarettes a lot. The likelihood of a "no association" result is very small, indeed a lot smaller if cigarettes are hazardous, than such a result if cigarettes are safe. Therefore, a negative result would weaken one's degree of confidence a lot. In the language of laboratory tests such evidence is both sensitive and specific.)

Predominantly strengthening: If EMFs were indeed hazardous this type of evidence doesn't have the quality or power to detect anything consistently but a very large effect. Thus, it is not very likely to give a positive result if the effect is small, but it is still more likely to be positive than would be the case if EMFs were not hazardous. Therefore, a positive result still can strengthen one's degree of confidence quite a bit. But the likelihood of a negative result is fairly large even if EMFs were hazardous, yet not quite so large as the likelihood of a negative results would be if EMFs were not hazardous. Therefore, a negative result weakens one's degree of confidence but only slightly. In the language of laboratory tests such evidence is specific, but not sensitive. The accumulation of studies of the predominantly strengthening type can eventually weaken the degree of confidence, and it can weaken the confidence that an easily detectable large effect is present. (An example from another domain would be studies of "secondhand" smoke. Here we have much more difficulty figuring out how much exposure people get. The expected effect is small compared to the ability of affordable studies to detect it. Therefore, even if secondhand smoke is hazardous we don't have a large likelihood of picking up the effect (although the likelihood is larger than would be the case if secondhand smoke were safe). So a positive result strengthens the degree of confidence. On the other hand, the likelihood of a negative result if second hand smoke is hazardous is pretty large and the likelihood of a negative result if secondhand smoke is safe is only slightly larger, so a negative result weakens the degree of confidence only slightly

Predominantly weakening: This type of evidence gives lots of false positive results, so the likelihood of a positive result when large is only slightly greater if EMFs were hazardous than if they were not. So a positive result doesn't change one's degree of confidence much. On the other hand, a negative result is relatively much less likely if EMFs were hazardous than if they were safe. Therefore, a negative result weakens one's degree of confidence considerably. Such evidence is sensitive but not specific.

Often a little reflection about a class of evidence or a particular study can give a good indication of whether a positive result will be as convincing as a negative result. It is a common tendency to assume that all evidence is of the "strengthening or weakening" variety. But this is not always the case.

Example of a Qualitative Bayesian Argument

One can illustrate the form of argumentation which we are advocating by applying it to the case of thalidomide. A series of babies without arms or legs was born to women who had taken thalidomide in early pregnancy. What evidence was available at the time on molecular structure and function, metabolic knowledge, animal tests and epidemiology?

The likelihood that a small epidemic of specific birth defects would appear after the introduction of thalidomide is quite a bit larger if thalidomide is hazardous than if it is safe, So one's degree of confidence of hazard increases quite a bit after reviewing the epidemic. This is particularly so when one notes that the medication was taken at the vulnerable time of development of the fetal arms and legs.

Examining the molecular structure of the agent did not suggest a mechanism for a hazard, but the likelihood of having that kind of explanation even if it were hazardous is small, though relatively larger than if the agent were safe. If one had a theory, it would boost one's degree of confidence, but the absence of theoretical mechanism doesn't pull down one's degree of confidence much.

Animal studies did not show thalidomide to cause birth defects at first. But the likelihood that something that causes birth defects in humans will do so in any given species of rodent is not very high, though higher than would be the case if the agent didn't cause birth defects in humans. So once again this stream of evidence can strengthen one's degree of confidence if one gets a positive result, but doesn't pull it down much if one gets a negative result.

What is the net result? Before one heard about the epidemic, one's initial degree of confidence that Thalidomide would cause birth defects was quite small ("very unlikely to cause" birth defects). That is because there are many medicines that are taken during pregnancy and only a tiny minority have ended up causing birth defects. The lack of mechanistic reasons and the negative animal study pulls that degree of confidence down, but not very much. The coherent epidemiological findings with big effects are relatively much more likely if Thalidomide is a hazard than if it is safe, and that pulls the degree of confidence up much more than the other streams of information pulled it down. So one ends up with a "highly probable" to "virtually certain" degree of confidence that Thalidomide causes birth defects.

The preceding argument did not use probability numbers, but followed steps of reasoning that are analogous to those that would be used in probabilistic reasoning. This is the kind of qualitative argument we propose to use. We believe that this will make more transparent the thought process linking the pattern of evidence and our subsequent degree of confidence about causality. If there is a stream of evidence in which the base is too sparse or biased in unpredictable ways or contradictory, then the likelihood of that pattern if EMFs were hazardous is similar to the likelihood if EMFs were safe, so the relative likelihood is not informative and will not influence the degree of confidence much.

How to Form an Initial Degree of Confidence

Considering what we know about physics, biophysical argument and general biology, what initial degree of confidence should we have of a causal explanation for the kind of barely detectable epidemiological associations compatible with the body of evidence which has now accumulated for certain diseases?

1) If there had been an anatomic structure for detecting residential level EMFs so that biological effects from them were biophysically explainable, what should our degree of confidence about pathological effects have been before seeing mechanistic, whole-animal pathology, and human evidence?

2) Should we have started out assuming that powerline magnetic or electric fields were as likely to be beneficial as harmful?

3) Should the proportion of chemicals and physical agents with hazardous properties at ambient levels influence our initial confidence of an EMF hazard?

4) How much should the initial confidence in item 1 be pulled down, given that we know of no such structure and there are biophysical arguments that combine physics and simple models of cells and tissues to suggest that residential and occupation EMFs should not be detectable and therefore should not produce either physiological or pathological change,? Do these theoretical arguments have the same strength that thermodynamic arguments and assumptions about friction in machines have about the impossibility of perpetual motion machines?

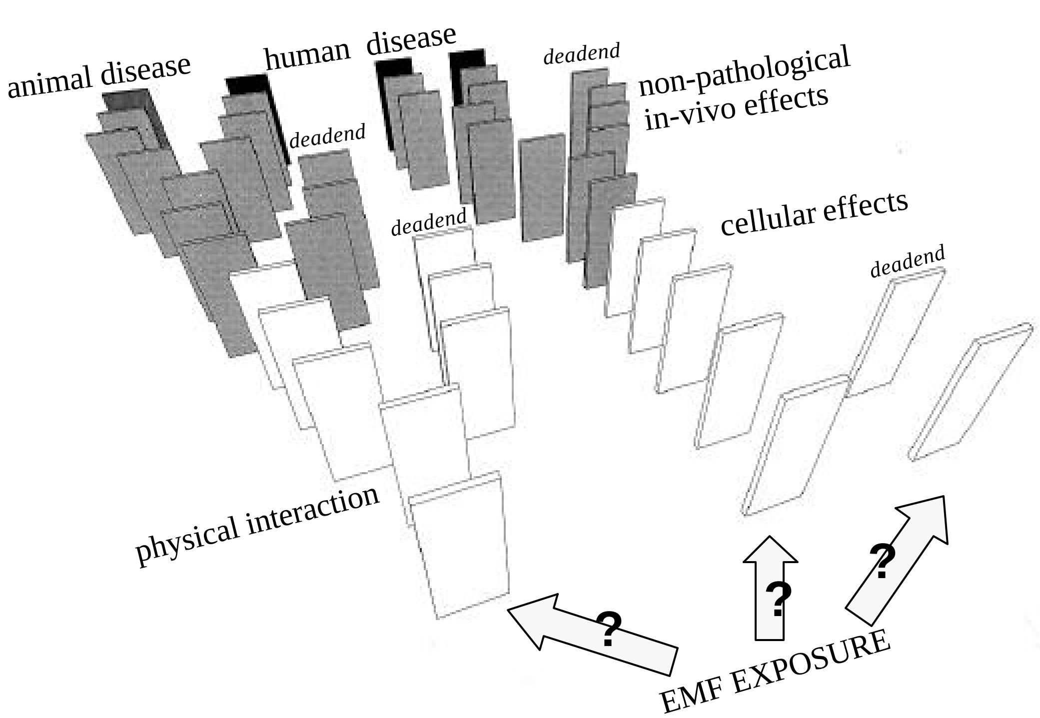
We view the biophysical theoretical stream when combined with general biological knowledge as being related primarily to our initial degree of confidence (although biophysics may also be relevant when discussing dose response results) and the mechanistic, whole-animal and epidemiological streams of evidence as being available to update that confidence.

When the time comes for eliciting our reviewers' degree of confidence, we will also consider in a qualitative way, using everyday English, the initial degree of confidence.

Coherent Evidence from Different Levels of Biological Organization

It is usually accepted that our degree of confidence about a causal relationship is increased considerably if we have evidence from several levels of biological organization. If we could show that EMFs produce molecular, cellular, physiological and pathological changes at several levels of organization, we would have more confidence that it was a hazard than if we had only epidemiological evidence. We should be more explicit about why this is so and should discuss if this combined pattern of evidence is likely to be "predominantly strengthening" or "strengthening or weakening" in nature.

Figure 1. Dominoes representing steps in a mechanism linking EMF exposure to disease.



Epidemiologists see only the black dominoes relating to humans, toxicologists only the black ones relating to laboratory animals, *in vivo* experimenters only the light gray, and cell biologists only the white.

The possible chain of events that may link EMF exposure to an observable pathological effect is schematically illustrated by the array of dominoes in Figure 1. Here the white dominoes represent effects at the cellular level, the light gray dominoes represent effects (not necessarily pathological) on living tissues and the black dominoes represent pathological effects in humans or animals.

When epidemiologists see one or more of the black human disease dominoes fall, they ask scientists in other disciplines to explain the reason for it. Exposure to EMF may or may not cause one or more of the first row of white dominoes to fall, and cell biologists may observe this. Whether this is connected to what epidemiologists have observed depends on the arrangements of the dominoes behind the black human disease dominoes. Some falling dominoes may cause only a few further dominoes to fall before the series comes to a dead end, such as an effect at the cellular or the tissue level that is well tolerated by the organism and does not result in an adverse health effect. For example, see the two white dominoes at the far right of the illustration.

The same domino may cause two parallel rows of dominoes to start falling. One of these may come to a dead end (as occurs for the white dominoes in the large left branch of the illustration) while the other may result in an observable disease in both animals and humans. This would correspond to the terminal branches to the left of the illustration. Sometimes other causal chains must be operating simultaneously for a chain of event to cause disease. This would be represented by converging lines of dominoes, which only together can topple a final black domino that sits beyond the function of the several lines.

Before we say anything further about the domino metaphor, it is wise to point out how it differs from reality. Most biological processes are not a linear progression of events. In real life there is redundancy and feedback loops producing complex systems that defy simple intuition. EMFs could cause a different physical induction mechanism within different intensity ranges, each with more than a one-cell physiological consequence. Many might not lead to any pathology at all. Thus, it may be possible through study of mechanistic research literature to construct a plausible story in which EMFs lead domino-like to human pathology, but that story is not guaranteed to be the truth. With this caveat let us proceed. How would scientists in the several lines of research view the chain of events represented by the depicted row of dominoes?

Epidemiology can see that the first domino fell (exposure to EMFs occurred) through exposure assessment. It should be noted, however, that measurement of exposure in epidemiology is often done after the fact and consequently is usually less precise than in experiments. Epidemiology also observes whether the last domino falls subsequently (disease occurred). In some cases, epidemiologists may measure intermediate steps, for example, effects of EMF exposure on production of the hormone melatonin.

Whole-animal bioassays provide a similar view of the first and last events. They differ in that the investigator causes the first domino to fall by exposing animals to EMFs and has complete control over this first step. Consequently, the whole-animal bioassays can quantify the exposure much more precisely. They also differ from epidemiology in that the investigator may expose animals to higher levels of EMFs than encountered in the environment. When considering results from animal studies, it is important then to consider how to scale results from smaller animals to larger humans.

In mechanistic studies, the investigator also causes the first domino to fall and then notices whether or not some intermediate domino (a biochemical or physiological step on the way to disease) falls. These studies do not provide much insight about the second domino (representing that first molecular reaction to EMFs) or the fate of other intermediate dominoes lying either upstream or downstream of the domino under investigation. One can assume, however, that if a step late in the series occurs earlier steps must also have occurred.

Biophysics concentrates on the second domino, the first biological response to EMFs. In the domain of the bioeffects of noise and ionizing radiation this kind of understanding has given some insight into exposure-response relationships and increased the degree of confidence that these agents can cause biological effects.

One's degree of confidence about causation increases if one can experimentally push the first domino oneself and see many of the intervening dominoes fall against each other on the way to the last domino. We can rarely document the entire causal process in humans or in experimental animals. For example, our understanding of the steps that lead from cigarette smoke exposure to lung or other cancer did not derive from experiments where each step was observed in humans or beagles. Rather, the evidence was pieced together from many different studies.

A series of experiments can document different segments of the hypothesized process in different organisms. If one had evidence of the "physical induction" mechanism and a series of physiological and pathological mechanisms from mechanistic and whole-animal experiments, it would increase our confidence that the EMFs cause a disease. However, if one had strong epidemiological evidence, one's degree of confidence may already be quite high, and one may have less need for increased confidence from mechanistic studies.

There are other reasons that a composite of experimental evidence about the chain of events leading to a disease tends for better or worse to increase the degree of confidence of most scientists that an epidemiological association is causal. First is the principle of "Ockham's razor." William of Ockham, a 14th century scientist and theologian, recommended use of the simplest theory to explain a finding.[6] If one had a glimpse of many intermediate dominoes falling on the way to the last one it would seem unreasonable to postulate a number of *rows* of dominoes independently causing the intermediate dominoes to fall in the correct sequence.

It seems to be generally true that scientists believe a simple and elegant explanation more than a complex *post hoc* theory. This heuristic tool, to which we are sympathetic, is however just that, a modeling tool. Sometimes the truth is complex. Second, observing the effect of an agent on many intermediate dominoes increases one's degree of confidence in that it helps to rule out methodological bias due to confounding as explanations for at least those steps. Third, scientists tend most to believe evidence from their own disciplines once it has passed their detailed criticism**.**

Aside from building credibility for the causal theory, mechanistic information can increase the precision of our predictions about how exposures are related to disease. For example, scientists' understanding of the molecular events in DNA that result from the exposure to ionizing radiation provided some rationale for a no-threshold dose-response model. But what happens to one's degree of confidence of causation if, as is usual, there is little or no understanding of the mechanistic pathway between an agent like EMFs and disease? The known human carcinogens with complete, detailed mechanistic explanations is low. Thus, if in fact EMFs are carcinogenic, the likelihood of complete mechanistic evidence by this time is low. The likelihood of convincing mechanistic evidence of carcinogenicity if EMFs are not a carcinogen is even lower. Complete mechanistic evidence is of the "predominantly strengthening" type. A positive result would increase our degree of confidence a lot, but negative evidence would not decrease it very much. If all known mechanistic pathways toward disease were shown to be unaffected by EMFs and its relevant attributes, one's degree of confidence would be repeatedly pulled down, always to a slight degree. If we knew of many such unaffected mechanisms, the cumulative effect could start to pull our confidence down substantially. It is important to remember that although we know a lot mechanistically, what we don't know is vastly larger, so the percent of possible mechanisms shown not to be effected by EMFs is necessarily small. The more general inference applies also to whether the mechanism is on a plausible path to disease. It would be more convincing if the mechanism was directly relevant to humans. A mechanism that could produce some kind of adverse effect is more likely relevant than one that results only in physiological adjustments.

The order in which we have listed types of evidence (human, *in vivo, in vitro*, mechanistic) is not random. From the public health standpoint, they go from the more to the less relevant. This is not equivalent to saying from the less to the more important. A biophysicist would order them differently, with good reason, since only when observation has been explained by theory can one claim to fully understand a scientific phenomenon. However, the purpose of this evaluation is more limited and pragmatic. Even if we could build a theoretical model that could perfectly explain how low level environmental magnetic fields are perceived by living organisms (notwithstanding a very low signal-to-noise ratio), we would still not know whether these fields pose a risk to human health. We would still need to show:

- that these fields, as well as being perceived by living cells, alter normal biological processes, including the physiology of the cell or the whole animal

- that these processes lead to adverse effects
- that these adverse effects are part of the causal chain leading to the disease we are considering.

If we could establish that the epidemiological evidence is completely convincing, we would not need to evaluate the previous areas of research to conclude that environmental EMFs pose risks that may warrant action. For example, the very strong association of Reyes syndrome with aspirin use in children has strongly increased our confidence of a hazard and has compelled warning labels and changes in pediatric practice, even though whole-animal and mechanistic evidence provide no support.

If we could prove beyond doubt the association with EMF of any of those steps we need not prove the preceding ones. In the more likely case that the evidence for each of these steps falls short of being conclusive, we will regard it as increasing the plausibility of the evidence supporting the steps that follow and decreasing the weight of the evidence running contrary to the plausibility of the steps that precede it.

From this it follows that for the purpose of these guidelines the weight given to each stream of evidence will depend on other types of evidence. The weight will be dynamic, depending not only on the intrinsic merit of the scientific discipline and of the sensitivity and specificity of the studies, but also on the need that the risk assessor has for that evidence in the context of the hierarchy outlined above. An assessment carried out when only weak epidemiological evidence is available may need to place more weight on mechanistic evidence (even if this carried a fairly high probability of false negatives) than if the epidemiological evidence were strong. However, should new evidence become available that, for example, substantially reduces the likelihood of confounding in the epidemiological studies, one would need to reassess the relative importance of the mechanistic evidence.

Whole-Animal Bioassays

Our degree of confidence as to whether EMFs cause disease in humans will be influenced by the relative likelihood conveyed by the pattern of evidence from whole-animal experimentation. This is tempered by the fact that different species of rodents react differently to some carcinogens and that, at least for many years, agents such as tobacco, arsenic and benzene, while causing cancer in humans, had no demonstrable carcinogenic effect in the species of animals tested. At this point, all recognized human carcinogens create cancer in at least some animal species, although not always in the same organs.[7] Some have argued that animal bioassays give as many as 50% false positives,[8] but others put this closer to 10%.[9]

EMFs are different from chemical agents and some other physical agents in at least two ways. One is that we are not certain what attribute of EMFs may cause effects. So, unlike chemical bioassays, where it is clear that the correct agent is being tested, in bioassays for EMFs we need to consider the attribute of EMFs used in the experiments. The second difference is that we are not certain that the risk from exposure to EMFs continues to increase from lower doses to higher doses. This is different from chemical carcinogens, where it is assumed that higher doses will cause higher risks and consequently that higher dose experiments will be useful in detecting effects even if small numbers of animals are used. It will be important to assess whether the high dose/large effect assumption should be carried over from chemical and ionizing radiation studies into non-ionizing radiation studies. If not, the traditional bioassays may not have the power that they do in the domain of chemical carcinogenesis.

Moreover, it is difficult to extrapolate between the exposure used in animal experiments and the environmental levels. We do not know whether the time-weighted average is the true “exposure metric” (see glossary). If repeated short exposure to elevated fields (conceivably strongly correlated to an abnormally high TWA) were the risk factor, the field used in animal experiments would not be orders of magnitude higher. If induced currents were a link in the interaction mechanism, allowance should be made for the small size of rodents, which results in smaller induced currents for a given field.

Biophysical Arguments

Usually, theory is built on the basis of observation and used to predict other observations. If evidence runs counter to these predictions we are compelled to question the evidence. However, if this stands up to scrutiny, the degree of confidence in the theoretical prediction falls. There are no situations in modern science in which theory takes precedence over observation.

Some scientists claim that all evidence of EMF health effects must be due to artifact. This claim is based on purely theoretical considerations: living organisms have a relatively high level of random electrical signals due to endogenous electrical currents and the Brownian motion of electrical charges. The weak environmental fields would not be perceived above this "noisy" environment. They have supported their point of view with two different approaches. The first consists of calculating the minimum signal strength that can be detected above this noise. The other approach is to derive, from the known necessary characteristics, the description of an organ capable of detecting the low environmental EMF levels and then to point out that such a detector would need to be so large that it would have been identified by now.

Biophysical arguments applied to EMFs rely on the laws of physics applied to simplified models of molecules. Our evaluation must address the question whether our experience with these combined models allow us the kind of confidence about predictions of effects of EMFs on biological systems that we have about perpetual motion machines. One can easily argue that perpetual motion could only exist if friction could be totally eliminated, because according to the second law of thermodynamics the energy dissipated by friction as heat cannot be recuperated. Since nobody questions the second law of thermodynamics or the inevitability of friction, perpetual motion is universally acknowledged to be unachievable.

However, in practice we all accept perpetual motion as a fact of life. Even though we know that the motion of planets and stars will eventually stop and the universe as we know it will eventually end as a consequence of the second law of thermodynamics, this does not prevent us from behaving as if it were eternal, since it is eternal compared to the time frame by which we live our lives. Therefore, both in the EMF and in the perpetual motion situations the problem is not one of possibility or impossibility, but one of realistic limits. The theoretical limits placed on EMF effects are only credible if the context in which they are derived is realistic. This is were the debate between proponents and opponents of the so-called "impossibility argument" really hinges. We see these arguments, if relevant at all, as being relevant to our initial degree of confidence of an EMF hazard effect prior to considering the pattern of results in the other streams of evidence.

The relationship between mechanistic evidence and biophysical arguments is quite important. Results that document a portion of a plausible mechanism, if convincing, could cause evaluators to give less weight to biophysical theory by showing bioactivity where the theory predicted none. Theory would then cease to influence the judgments about the credibility of other experiments or the epidemiology. Incomplete mechanistic research results, if convincing, could build our degree of confidence that EMFs were bioactive at environmental levels. This in turn would build the credibility of epidemiological associations with environmental level EMFs. The degree of confidence about environmental epidemiology would be increased more by bioactivity at environmental levels of EMFs than by bioactivity at levels far above those found in the environment. The confidence is increased because if the agent is active at this level it could also be harmful. If it is not active at this level then it cannot be harmful.

What if the agent is bioactive despite biophysical predictions, but at a level far above those found in environmental settings? This might contradict predictions of biophysics that no effects occur at this level, but it is not clear that the bioactivity is relevant to environmental exposures.

Evidentiary Tests for Causality in Epidemiology

Bradford Hill, a well-respected statistician, proposed nine attributes to which epidemiological associations can be compared to consider whether they are likely to reflect cause and effect.[10] These were strength, consistency between studies, specificity, temporality, biological gradient, plausibility, coherence, experimental evidence and analogy. Rothman and Greenland describe the limitations of these criteria.[11] Consider the criterion of "specificity," the notion that a single agent causes a disease. Many agents cause more than one disease. Nowadays, we recognize that even infectious organisms like tuberculosis and syphilis can cause pathology in different organs with vastly different symptoms. Smoking causes a variety of cancers in organs as disparate as the lung and the bladder and causes heart disease and chronic lung disease as well. It could be argued that a physical agent would be less specific in effect than a biological or chemical agent. The structured questions relevant to epidemiological data capture all of Hill's questions, but frame them so as to encourage a graduated kind of answer. What they do not do is generate a checklist of Hill questions and add up the yes and no answers.

State of the Science

Evaluating the rate of progress in a scientific field and predicting which approaches are likely to yield results is not easy. Peer-review groups and funding agencies do their best to pick promising lines of research. The case-by-case evaluation of investigator-initiated research does not yield an overview of the research field to policy makers. California Department of Health Services staff will consult with the World Health Organization and Electric Power Research Institute to assess what research is in the pipeline and what areas are not being researched at present. DHS staff will provide an opinion based on a decade of following the research field, on areas of research if any, which might produce useful information. They will also provide pro and con and summary arguments to justify their opinions and their estimates of the duration of any needed research effort before positive or negative results become probable. They will provide a range of estimates. The policy analysis of our contractors helps spell out the implications of these estimates for research policy.

V. Summary of the Ideal Approach

The approach will rely on reviewing and extracting information from existing analyses and key studies and will start with the detailed reviews compiled by the National Institute of Environmental Health Sciences. Our goal is to provide a useful and informative interpretation of the evidence rather than an extensive listing of factual evidence. When new studies are crucial to influencing the degree of confidence one way or the other we will summarize them in somewhat more detail. We will discuss the issues related to the many attributes of EMFs and the ways that they can be measured. For the analysis, we will select key exposure metrics as the focus. We will explicitly identify disease outcomes to be included.

To assure systematic attention we will use a structured set of questions for each stream of evidence. We will make our case as to whether this stream is " uninformative," " strengthening or weakening," "predominantly strengthening" or "predominantly weakening." We will use the device of pro and con and summary arguments to assure that we are not ignoring evidence or arguments and to make our thought process open to public comment and challenge. In these arguments we will contrast the likelihood of finding this pattern of evidence if EMFs were hazardous with the likelihood if EMFs were not hazardous.

After the EMF project team summarizes the evidence and prepares the pro and con and summary arguments, other environmental scientists in DHS will be asked to review the original literature and critique the summary and the pro and con arguments. The core team and critics will then meet to review the revised pro and con and summary arguments and the consideration of what the initial degree of confidence should have been. Everyone will provide an anonymously written "initial best estimate," an upper bound and a lower bound of the degree of confidence number. Those with outlying values will anonymously defend their positions in writing, and the group will vote again. Graphs of the distribution of best estimates will be presented to provide decision-makers information about the range of degree of confidence among the responsible DHS scientists who have been asked by the PUC to make this determination. The distributions will be summarized using narrative phrases.

Here are examples of possible results:

> None of those voting had an upper bound degree of confidence that EMFs caused x that exceeded "very improbable to cause." For the purposes of the policy analysis the Department would recommend using confidence numbers between 0 and .09, although 90% of the DHS scientists had best estimates which clustered around 0.001.
>
> All of those voting had a lower bound degree of confidence that EMFs caused y which ranged between "probable, more likely than not" to "highly probable that it is a cause." For the purposes of the policy analysis the Department would recommend using confidence numbers between 0.51 and 0.97, although 90% of the best estimates of DHS scientists clustered tightly around 0.90.
>
> All of those voting had a wide range between their upper and lower bound degrees of confidence and their best estimates varied greatly from person to person because of the small size of the evidentiary base and its contradictory pattern and poor quality. The Department scientists were unable to pinpoint a defensible degree of confidence for use in the policy analysis.

We will attempt to estimate the magnitude of relationships. We will also consider the individual lifetime theoretical risk and the attributable population burden. We will discuss the state of the science and the likelihood and imminence of scientific breakthroughs that might change the results.

Possible Simplifications

If restrictions in time and manpower make it impossible to carry out all the above steps for all diseases of interest we will focus on those diseases with the most information and the highest incidence. We may not assess the state of the science for all streams of evidence or all diseases.

PART TWO: OUTLINE AND SPECIFIC GUIDANCE FOR THE RISK EVALUATION

This second part of the Risk Evaluation Guidelines provides guidance to the California Department of Health Services staff and consultants who will be conducting the risk evaluation following the principles and approaches described in the first part.

The evaluation will consider all reports published in the peer-reviewed literature by March 31, 2000. Studies with limitations (e.g., no quantitative exposure assessment) or flaws (e.g., selection bias) will be evaluated in the light of such limitations, and an effort will be made to investigate their possible consequences. Data generated by the California EMF Program will be evaluated after external peer review or acceptance for publication. If any of the following crucial epidemiological studies become available after acceptance in a peer-reviewed journal we will consider them and integrate them into the document by June 30, 2000.

> British collaborative childhood leukemia study
>
> Seattle breast cancer study
>
> USC breast cancer study

Kaiser Foundation Research Institute miscarriage study

Pooled analysis of childhood leukemia studies by Greenland, Shepard, et al.

Manuscripts presented at the California EMF Program Epidemiology Workshop (Berkeley, January, 1999), even if unpublished, will be regarded as briefing documents for the evaluators, since the stated goal of that workshop was to assist the DHS evaluators in their task. If relied on heavily they will be included as appendices to the report and will be open to public comment.

The evaluation will be conducted by a team of scientists from DHS and the Core EMF Program scientists representing several disciplines. Outside consultants will also be involved in the preliminary summary of the pattern of evidence, but not in developing the pro and con arguments.

I. Disease Endpoints and Exposures of Interest

A. Health Endpoints That Will be Considered

We will consider the same group of diseases considered by the National Institute of Environmental Health Sciences (NIEHS) Working Group report.[1] These are diseases for which there is some epidemiological evidence of an association.

The NIEHS working group report discussed the following diseases:

- childhood leukemia
- chronic lymphocytic leukemia
- acute myeloid leukemia
- brain cancer
- breast cancer (male and female)
- central nervous system cancers
- childhood central nervous system cancers
- childhood lymphoma
- reproductive health (mother and father exposure)
- Alzheimer's
- amyotrophic lateral sclerosis and other motor neuron diseases
- suicide and depression
- cardiovascular diseases
- electrical sensitivity

We are focusing on diseases and not on harder-to-identify functional endpoints such as sleep disorders, learning difficulties, etc., except as they might be relevant to disease causation.

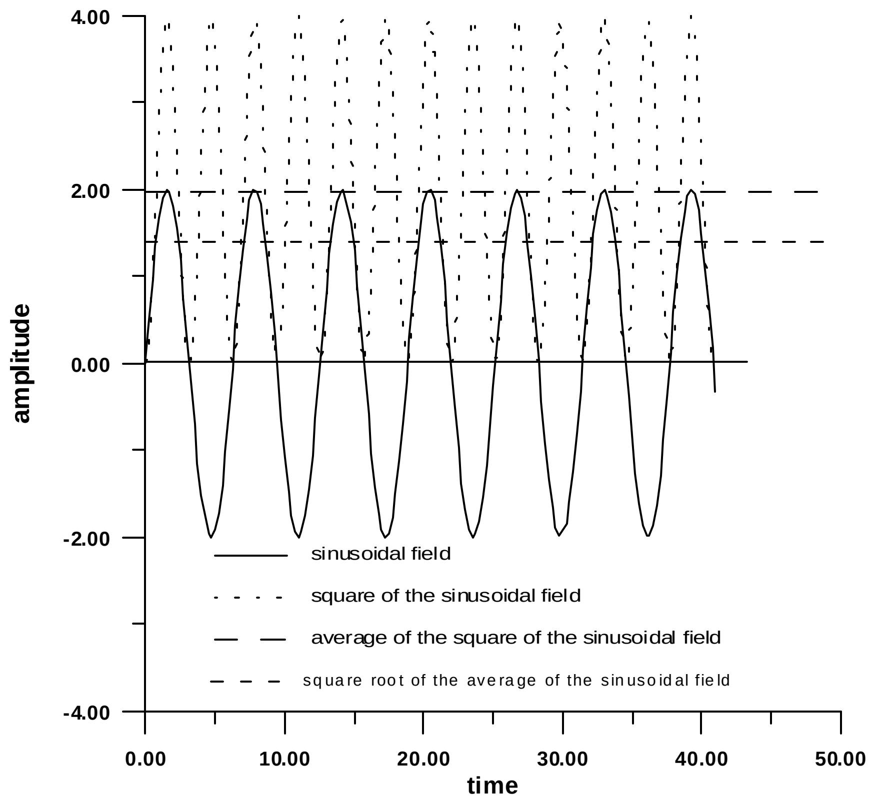
B. EMFs: Types of Fields, Levels and Frequencies Included in the Risk Evaluation

The risk evaluation will focus on certain kinds of electric and magnetic fields. This section provides a brief summary of what creates and defines these fields.

These guidelines are concerned only with fields resulting from the generation, transmission, distribution and use of electric power. They do not include other kinds of fields, such as those associated with cell phones. The guidelines described in this document are aimed at effects at intensities other than those due to induced currents, which are well understood.

Since the electric and magnetic fields from power lines oscillate symmetrically around a zero value many times a second, their magnitude cannot be measured by their average (since this is always zero). The magnitude of one of these 1/60th second cycles can be expressed either as the absolute distance between the peak at the top of the cycle and the peak at the bottom of the cycle (peak-to-peak) or by squaring each of the instantaneous values, taking the average of these squared values and then taking the square root of this average. The later is called the “root mean square” (rms) value of that cycle. Most instrumentation aims at capturing the rms value. The idea is illustrated in the figure below.

Figure 2. The root mean square (rms) is equal to peak-to-peak value divided by 2.82 (i.e., 2√2). In the example below the peak-to-peak value is 4 and the rms is 1.41.



There is no biological reason to use any specific measure of the magnetic field in epidemiological studies. Although most instruments measure the rms, at least one study measured peak-to-peak exposure.[12] Categorical exposure assessments (“wire code,” job description) are correlated in different degrees to several aspects of the EMF.

Literally, the term “time-weighted average” (TWA) refers to the practice of measuring rms exposure in different environments and averaging the results after weighting them according to the time the subject spent in each environment. In reality, only one study[13] followed this approach. In other studies, when exposure was inferred through measurements or calculations (as opposed to qualitative means such as wire-coding), measurements were averaged (with no weighting, since none was required) over the duration of the measurement in the residence or when doing a work task or, in the case of calculations, over one year. In this document we will use the term TWA to refer to a metric that captures the strength of the field averaged over a period of time sufficiently long to characterize chronic rather than accidental exposure. The evaluators may decide that the evidence is sufficient to adopt a more specific definition.

The broad definition given above does not allow differentiation based on other aspects of the field. This situation is analogous to many in observational epidemiological research. For example, in diet studies, one can correlate the consumption of red meat to adverse health effects without distinguishing between the various attributes of red meat. The assessors will be asked to decide whether the evidence is sufficient to differentiate between the TWA fields produced by powerlines or appliances.

The evaluation will rely heavily for factual matters of exposure on the exposition and summary in the NIEHS Working Group Report of 1998 and individual studies where needed.

Valberg et al. classify the aspects of the EMF mixture into four categories:[14]

- frequency (harmonics, transients etc.) (see glossary)
- intensity and timing (intensity of the various frequencies over a longer time scale)
- spatial characteristics (polarization, uniformity over space)
- combinations (certain combinations of alternating and static fields, electric and magnetic fields).

To address the relevance of aspects of the field other than the TWA to any bioactivity or pathogenicity we will ask the following questions:

> What attributes of the EMF mixture that have been hypothesized to be bioactive or pathogenic are correlated with the TWA magnetic field strength?
>
> Of this smaller subset of attributes what is the evidence that would suggest bioactivity or pathogenicity of this aspect in residential or occupational settings?
>
> On the basis of this assess the plausibility of these attributes of the EMF mixture as candidates to explain observed epidemiological associations.

C. Distribution of Exposures in the Population

We will use the 24-hour TWA 30-300 Hz for men and women from Zaffanella's thousand-person study as an approximation of exposure distribution in California.[15] The personal exposure of small children will be derived from McBride's control group.[16] The prevalence of various wire codes in southern California will be estimated from the control group of London et al.[17] and the prevalence of wire codes in suburban Northern California from Lee et al.[18] (We will take an average of the last two weighted by the size of the populations north and south of San Luis Obispo.) We will estimate the prevalence of persons in electrical occupations from the 1990 census and of utility employees from data from the PUC.

II. Examining Physical Theory and Experimental Evidence

In this part of the evaluation we propose to systematically review three types of evidence regarding potential health effects associated with exposure to EMFs: biophysical arguments, experiments focusing on mechanisms, and animal studies looking at exposure and pathological outcomes.

A. Biophysical Models and Physical Arguments

Recognizing that cells obey the laws of physics and chemistry, physicists have developed biophysical models to explain how EMFs and living systems interact. The biological parts of these models are simplified representations of reality. These models indicate that the signal-to-noise ratio is too low for environmental EMFs to be detected in the noisy electric environment of living organisms. The proponents of these arguments claim that they are soundly based on experiments and provide a secure limit to the possible. Other scientists argue that this limit is a consequence of the models' limitations and point out that, as models have become more refined the predicted minimum detectable level has been revised downward.

1) Structured review

a) Review and explain the predictions of the arguments that some scientists make about how EMFs may or may not affect biological systems based on biophysical models and physical theory. Are the arguments based solely on theoretical physics or do they also encompass assumptions about biological systems? If so, what are these assumptions?

b) Can these theories and predictions best be viewed as uninformative, predominantly strengthening, predominantly weakening, or both? What are the implications for their interpretation?

c) Evaluate empirical results relevant to the physical arguments for environmental levels of EMFs and for environmental levels higher than average, but lower than those at which effects are clearly explained by the well-understood mechanism of induced currents. Consider magnetic fields expressed as the rms field strength and as the square of this.

Identify any aspects of the physical arguments that can be tested empirically, including predicted results.

Have any assumptions on which models have been built been tested? With what results? Do they support or contradict the theories?

Do any experiments support or contradict the predictions based on biophysical models? With what results? Do they support or contradict the theories? Specifically, do experiments show physiological or other effects at levels of EMF exposure where biophysical arguments predict that there should be none?

d) Consider how the physical arguments have evolved over time.

Have predicted thresholds for effects of EMFs on living systems been consistent or have predictions been changed often to incorporate new findings?

Have the model assumptions required adaptation to reflect empirical findings?

e) Assess implications of the physical arguments for understanding of the relationship between exposure and response.

Are the data together with the modeling arguments strong enough to derive expectations for the magnitudes of health effects for relevant levels of exposure? Can the argument be used in interpreting exposure-response information?

Do the arguments have implications for extrapolating results observed or conjectured at one dose level to another?

Could they inform experimental or epidemiological results in making an overall determination and/or aid in the definition of exposure metrics, design of experimental protocols, or understanding of expectations for dose-response relationships?

2) State of the science

a) To what extent can the biophysical theoretical analysis inform work on the biological mechanisms?

To assess this, evaluate the level of collaboration between proponents of physical arguments (physicist/modelers) and biologists conducting related experiments.

Is the proportion of publications on biophysical "impossibility" arguments that display active collaboration between physical theorists and biologists high, medium or low? Is this argument an example of one discipline criticizing another or a cooperative venture where serious efforts at joint clarification have been made? "Impossibility" arguments that result from prolonged serious physics/ biology collaboration should get more weight than those arising from a single discipline

b) Discuss the completeness and quality of research in this area and the prospects that future research would resolve outstanding questions. If the field has thoroughly researched relevant topics, this would suggest that findings should be given more weight. Have theories suggested experiments? Have experiments been pursued to their logical conclusion?

c) What future studies, if any, would be likely to provide useful results related to this topic? How soon could a breakthrough occur?

3) Pro and con arguments and resolution

Assess the evidence presented as physical arguments as a whole for or against a relationship between exposure to EMFs at environmental levels and at environmental levels higher than average but lower than those at which effects are clearly explained by the well-understood mechanism of induced currents and disease or biological effects that could lead to disease. Does the argument as a whole, including the biological assumptions made, the experimental evidence and the evolution of the argument, offer evidence for or against the existence of biological or health effects?

State the argument as it would be made to support the assertion that EMFs do not have biological effects at these levels.

State the argument as it would be made to support assertions that EMFs do have biological effects at these levels of exposure.

Fairly weigh the contrasting statements to give a judgment on whether or not EMFs cause biological or health effects, again for these levels of exposure.

What is the proper direction and magnitude of the effect of biophysical "impossibility" theories on one's initial degree of confidence that the range of EMFs from residential and occupational exposures could cause bio-effects or pathology?

B. Results of Mechanistic Studies and Biological Experiments

Studies of cell systems, tissues and other types of assays, along with animal and human studies focused on mechanisms, will be reviewed for evidence of molecular or cellular processes or other mechanisms that could relate exposure to EMFs to health effects.

1) Structured review

a) Identify the most useful reviews of biological experiments and mechanistic experiments. Review those most relevant.

Identify the particular studies that are most informative about mechanisms for any potential effects of EMFs. Include studies that have been replicated in two or more laboratories or that are considered to be of high quality even if not replicated.

Identify studies at relevant exposure levels that may be helpful for assessing whether EMF exposure causes adverse physiological effects of concern, identifying the causal pathways for producing those effects, and analyzing dosimetry options. The mechanisms and effects considered should include mechanisms relevant to carcinogenesis, directly through genetic damage (i. e. DNA breakage) or through signaling processes that may promote cancer development, as well as mechanisms relevant to non-cancer outcomes.

It may be that only a few examples would merit detailed analysis. The strategy is to examine in detail the most relevant biological evidence for effects.

Describe results of the studies identified as most informative.

b) Can these studies best be viewed as uninformative, predominantly strengthening, predominantly weakening, or both? What are the implications for their interpretation?

c) Does the mechanistic evidence rest on only one level of biological organization or is it supported by some combination of molecular, cellular, tissue, organ and whole-animal or human studies?

d) Is the proportion of studies showing a mechanistic effect of EMFs high, medium or low? Is the proportion of studies which have used EMF exposures that mimic the exposure to the EMF mixture in residential or occupational settings high medium or low? What are the frequently explored isolated aspects of the “mixture”?

Do the studies detect physical induction of responses to EMF exposure or do they detect biological responses?

Are physiological effects in cells or animals from exposures comparable to those found in environmental or occupational environments or are they from exposures at a much higher level? Consider where appropriate interspecies scaling factors for exposure.

Are any of these effects linked to causal chains leading to pathology in general? How convincing are the links?

Are any of these physiological effects linked to causal changes that lead to specific pathologies identified through epidemiological studies as relevant to EMF exposure? Do physiological feedback mechanisms or repair mechanisms compensate for or correct these changes? How convincing is the evidence?

Does any combination(s) of mechanistic findings appear to fit together into a coherent set of hypotheses that transcends more than one level of biological organization. If so, which? Describe the combinations or why none could be found. How convincing is the evidence?

Is the proportion of mechanistic studies using exposures which mimic the EMF mixture actually seen in residential and occupational environments high, medium or low? How does this affect the relevance of the findings?

2) State of the science

a) Do mechanistic results help interpret existing epidemiological results or suggest better ways for future studies to assess physiological measures of exposure or effect or to carry out exposure assessments?

b) Discuss the completeness and quality of research in this area relevant to hazard assessment and dose-response, considering the volume and content of publications and professional presentations so far as to whether there are promising leads which have not been followed up or inconsistencies which need to be resolved.

Based on the history of successes and failures of replication for different mechanistic hypotheses and measurement systems and the history of increasing complexity of mechanistic theories, what are the most pertinent experiments that could be performed to provide evidence of whether and if so how EMFs produce biological responses related to specific effects? How likely is it that these salient questions will be resolved to the satisfaction of most fair observers by further research in the next five, ten, or twenty years?

c) What future studies, if any, would be likely to provide useful results related to this topic? How soon could a breakthrough occur?

Consider whether given the efforts expended so far there has been a significant shortage or inconsistency in findings.

3) Pro and con arguments and resolution

Assess the evidence relating to mechanisms by which EMFs might have effects on living systems. Do the studies as a whole, considered across different levels of biological organization, offer evidence for or against the existence of health effects at relevant exposure levels? Consider in this judgment the replication or failure to replicate results, the extent of positive findings compared to the efforts expended, and the consistency, if any, across levels of biological organization and levels of exposure. Is the body of evidence strengthening or weakening, predominantly strengthening, predominantly weakening or uninformative?

a) Describe the biological evidence as it would be described to support the assertion that EMFs do not cause biological effects or effects that might lead to disease at relevant exposure levels.

b) Describe the biological evidence as it would be described to support assertions that EMFs do cause biological effects or effects that might lead to disease at relevant levels of exposure.

c) Fairly weigh the contrasting statements to give a judgment about the biological evidence and the weight to be attached to it at relevant levels of exposure. Explain this judgment.

d) Characterize the likelihood of the mechanistic study pattern of evidence if EMFs were indeed hazardous relative to the likelihood of this pattern of evidence if EMFs were not hazardous. Is this relative likelihood quite large, close to one, a small fraction? Which direction does this relative likelihood move your prior degree of confidence and by how much? A lot? A little?

e) Assess the implications of mechanistic findings for understanding dose-response relationships and for extrapolating results from one exposure level to another. Does the evidence provide any basis to select an exposure-response model?

C. Whole-Animal Studies Focused on Disease Outcomes

Scientists have studied the effects of EMFs on animals, particularly rodents. In this section, evaluators are to review the results of such studies.

1) Structured review

a) Identify the important studies of animals for consideration in this review. Summarize the animal studies, EMF attributes tested and levels, and outcome, including studies that consider EMFs as a cancer initiator or promoter, or reproductive or developmental hazard, or cause of other effect.

b) Can these studies best be viewed as uninformative, predominantly strengthening, predominantly weakening, or both? What are the implications for their interpretation? Consider whether bioassays at high EMF exposures have similar expected sensitivity and specificity as bioassays of chemicals at maximally tolerated doses. Do the animal studies provide evidence of a dose-response relationship with increasing response with increasing dose?

c) Discuss the applicability of animal studies. Discuss the power of animal studies and the issues associated with the need to test many animals to see the effect expected from epidemiological studies. Discuss the appropriateness of any extrapolation to lower doses from higher doses.

d) Discuss the sensitivity and specificity of the bioassays of one attribute of a mixture to predict the effects of the whole mixture.

e) With regard to the potential for carcinogenicity, what is the significance of bioassays of promotion, co-promotion and initiation of the process of carcinogenesis?

f) Is the proportion of whole-animal studies which have used EMF exposures that mimic the exposure to the EMF mixture in residential or occupational settings high, medium or low? What are the frequently explored isolated aspects of the mixture?

g) Do animal studies produce results that are incompatible with the predictions of current biophysical models?

2) State of the science

a) Discuss the completeness and quality of research in this area and the prospects that future research would resolve outstanding questions. If the field has thoroughly researched relevant topics, this would suggest that findings should be given more weight. Are there issues of study design that limit the applicability of results to date that could be corrected in future studies?

b) What future studies if any would be likely to provide useful results related to this topic? How soon could a breakthrough occur?

3) Pro and con arguments and resolution

Assess the evidence as a whole. Do the studies offer evidence for or against the existence of health effects at relevant exposure levels?

a) Describe the whole-animal assay evidence as it would be described to support the assertion that EMFs cause biological or pathological effects at relevant levels of exposure.

b) Describe the whole-animal assay evidence as it would be described to support the assertion that EMFs do not cause biological or pathological effects at relevant levels of exposure. Is this type of evidence strengthening and weakening, predominantly strengthening, predominantly weakening or uninformative?

c) Fairly weigh the contrasting statements to give a judgment about the evidence and an explanation.

d) Characterize the likelihood of the mechanistic research pattern of evidence if EMFs were indeed hazardous relative to the likelihood of this pattern of evidence if EMFs were not hazardous. Is this relative likelihood quite large, close to one, a small fraction? Which direction does this relative likelihood move your prior degree of confidence and by how much? A lot? A little?

e) Assess the implications of animal experiments for understanding dose-response relationships and for extrapolating results from one exposure level to another. Does the evidence provide any basis to select an exposure-response model?

III. Epidemiology Combined with Experimental and Physical Evidence for Disease Outcomes

A. Issues in Assessing Epidemiological Evidence across Diseases

Should the credibility or lack of credibility of occupational study results affect the credibility of residential study results and vice versa?

Should the credibility or lack of credibility of adult study results of a disease influence the credibility of childhood study results of the same or similar diseases?

Should the credibility or lack of credibility of results relating to one class of disease influence the credibility of results relating to another class of disease?

B. Insights from Mechanistic and Whole-Animal Studies

Can mechanistic studies be used to define more appropriate exposure metrics?

Has there been sufficient interaction between epidemiology and mechanistic studies? Is more effort at integration warranted?

Do the mechanistic observations provide insight into the observations or lack of observations of a relationship between exposure and disease response?

C. Epidemiological Evidence

1) Structured review

The following approach will be followed for each disease, or group of diseases, identified in the scope of review.

a) Compile information from epidemiological studies

Use authoritative compilations and reviews as the starting point for the evaluation. Identify further studies not considered in the compilations that should be considered in this evaluation. Include meta-analyses that provide useful and informative estimates of direction and magnitude of effects from analysis of multiple studies.

b) If it makes sense to group some diseases together for further consideration, identify which these are and the way they should be grouped.

c) Can these studies best be viewed as predominantly strengthening, predominantly weakening, or both? What are the implications for their interpretation?

d) Summarize results from compilations and additional information by disease or class of disease. For the diseases for which there has been considerable study, indicate the exposure setting and ranges of exposure. Describe:

the population studied

the exposure metrics or surrogates studied the results obtained, with any quantitative characterization presented and confidence intervals

e) Issues of study design and capacity. For each disease or class of diseases, discuss the following questions as they pertain to the body of evidence:

What is the direction and magnitude of bias (if any) introduced by the method used to select cases or controls (in case control studies)? If significant, could these problems be avoided in future studies?

What is the expected direction and magnitude of bias (if any) introduced by the method of measuring exposure in this series of studies? Could these be avoided in future studies

What is the expected direction and magnitude of bias (if any) introduced by any method of recalling exposure in this series of studies? Could these be avoided in future studies

Are there any well-recognized causes of the disease whose potential confounding effects were not dealt with in the design and analyses of enough of the studies so that they provide a likely alternative explanation for the associations or lack of associations seen? If so, what is the direction and magnitude of bias? Do we know of any risk factors for this disease?

Are there any weakly documented potential causes of the disease which were not dealt with in the design and analyses of enough of the studies to provide a likely alternative explanation for the association or lack of associations seen? If so, what is the direction and magnitude of the bias?

What kinds of studies could test the contributions of any of the two types of confounders discussed in above?

Consider an unspecified agent correlated with both the exposure surrogate measures and the disease. How strong would these correlations need to be for this agent to fully explain the observed association? How plausible is it that such an agent exists and remains unacknowledged as a risk factor?

Considering the imperfect correlation between exposure surrogates (e.g., wire-coding) and possible bioactive aspects of the EMF "mixture," how strong would the correlation with the disease need to be to result in the observed odds ratio? Is this plausible?

How likely is it that the cohorts chosen for studies could have had unique sensitivities to EMFs not representative of most people?

Can differences in dosimetry or exposure patterns plausibly explain differences in results between epidemiological studies?

f) Answer the following questions related to the capacity of studies to detect an effect of interest.

When taken together, what magnitude of effect would this series of studies have had the power to detect and with what resolution power?

Is there any biological evidence to expect an effect above or below this?

g) Answer the following questions related to the consistency of the studies.

Is there consistency or heterogeneity in direction or magnitude of effect between studies? Can we explain any heterogeneity across studies?

h) Specificity: If EMFs cause different variants of the same disease in different study locations, how does this affect your assessment of the evidence? If EMFs are associated with other diseases besides this one, how does this affect your evaluation of the evidence related to this disease.

i) Environmental justice: Is there any evidence that EMFs may particularly affect any identifiable segment of the population due to high exposures, heightened susceptibility, or other reasons?

j) Comparison to other risk factors: How does the apparent strength of association compare to that for other, more accepted risk factors for each disease or class of diseases?

k) Visibility: The increasing use of electricity and its ubiquity leads to a common sense expectation that EMF effects would be readily observable over the years and from highly electrified to less electrified areas. Explicitly assess whether epidemiological evidence (if any) for this disease would suggest that this is so.

l) For those diseases or class of diseases for which there is relevant evidence, characterize relationships between exposure and response in the body of evidence as follows:

Is there evidence of a dose-response relationship as measured by tests for trend?

Do studies that investigated subjects exposed to unusually high fields (e.g. electric welders, electric train engineers) report possible relative risks much higher than those reported in studies of populations comparing exposures in the 3-5 mG range to exposures in the 0-1 mG range? How much higher?

Is there evidence for or against a threshold of effect with regard to surrogates or measurements? If there is such evidence where is the threshold?

Is there any evidence for or against an upper plateau of effect with regard to surrogates or measurements? If there is such evidence where is the plateau?

Is there any evidence of an anomalous dose response relationship, such as a lower risk at the 95th compared to lower percentiles of exposure? How much lower?

Is there any epidemiological evidence of circadian or other biological windows of vulnerability in this disease? What is the magnitude of this effect modification?

Does this body of evidence provide any clue as to which EMF attribute is bioactive? For example, do 60 Hz studies show different results from 50 Hz studies? If so, how much?

Does this body of evidence provide clues as to the required duration of exposure or the interval between exposure and the appearance of disease? If so, what?

2) State of the science

Discuss the completeness and quality of the body of epidemiological research in this area and the potential that further replicating and predominantly strengthening or predominantly weakening research could contribute useful information to this analysis.

a) What could new epidemiological evidence contribute to this picture?

b) What new studies or compilations of studies are in the pipeline?

c) How imminent is any new information likely to be?

Could further epidemiological studies of this disease advance knowledge? If so what design features would be desirable?

3) Pro and con arguments for epidemiological evidence and for all evidence

a) For each disease describe the best reasonable argument that would be made from epidemiological evidence to assert that EMFs are a cause of the disease. Discuss relevant issues of effects of chance, confounding, misclassification, or other internal problems, as well as internal consistency of the studies and consistency across studies. Consider in addition your initial confidence (see questions in Part One) and evidence from other streams of evidence and use the weighting discussed previously to give for each disease the best reasonable argument, considering all the evidence, to assert that EMFs are a cause of the disease.

b) For each disease describe the best reasonable argument that would be made from epidemiological evidence and the other streams of evidence to assert that EMFs are not a cause of the disease. Discuss relevant issues of effects of chance, confounding, misclassification, or other internal problems, as well as internal consistency of the studies and consistency across studies.

c) Fairly weigh the contrasting arguments and the initial degree of confidence questions in Part One and give a balanced judgment of the degree of certainty that EMFs cause the disease at relevant levels of exposure

d) Provide a characterization of the confidence of this conclusion using categories of Table 1 and using the categories of WHO in Table 2.

e) Compare the weight of this evidence to other cases where epidemiological data were used to determine whether a compound was carcinogenic.

For diseases for which less information is available, a comparison should first be made to the cases with more information, and a judgment then made of whether it is appropriate to make a categorical and quantitative statement.

f) Assess the implications of epidemiological studies for understanding dose-response relationships and for extrapolating results from one exposure level to another. Does the evidence provide any basis to select an exposure-response model?

Table 2. WHO categories for classifying carcinogens by weight of evidence

	Description of Evidence	**IARC** International Agency for Research on Cancer **Classification**[19]
1.	Sufficient evidence from epidemiological studies.	1 carcinogenic to humans
2.	In exceptional cases less than sufficient evidence in humans, with sufficient evidence in animals and strong evidence in humans that the agent acts through a relevant mechanism of carcinogenicity.	1 carcinogenic to humans
3.	Limited evidence from epidemiological studies with sufficient evidence from animal studies.	2A probably carcinogenic to humans
4.	Sufficient evidence from animal studies with strongly supportive evidence from other relevant studies.	2A probably carcinogenic to humans
5.	Limited evidence from epidemiological studies with strong supporting data.	2A probably carcinogenic to humans

6. Sufficient evidence from animal studies.	2B possibly carcinogenic to humans
7. Limited evidence from animal studies with strongly supportive evidence from other relevant studies.	2B possibly carcinogenic to humans
8. Limited evidence from epidemiological studies with no or inadequate supporting data.	2B possibly carcinogenic to humans
9. Limited evidence from animal studies with no or inadequate supporting data.	3 not classifiable as to carcinogenicity to humans
10. Inadequate evidence from epidemiological, animal, or other relevant studies.	3 not classifiable as to carcinogenicity to humans
11. Sufficient evidence from animal studies with sufficient data to show these studies are not relevant to humans.	3 not classifiable as to carcinogenicity to humans
12. All available evidence suggests lack of carcinogenicity.	4 probably not carcinogenic to humans

IV. What the EMF Program's Decision Model Requires for Estimates of the Magnitude of Risk

One of the purposes for the evaluation is to develop estimates that can be used in a decision analysis model being developed in other parts of the EMF project to consider policy options. This model requires an estimate of the magnitude of effects for each disease. In this part of the assessment, evaluators will develop the estimates needed for the decision analysis. We will have to deal explicitly with whether information about dose response in one disease is relevant for another disease. We will review first any diseases where evidence is relevant to estimating the likely shapes of dose-response curves (if the association with EMFs was causal in nature). We will then consider the diseases for which some time-weighted average field strength (TWA) information was available and discuss whether there are dose-response curve types which are compatible with all the considered diseases and what the range of these are.

The decision analysis models have made some important assumptions about the way that exposure to EMFs is related to adverse health effects. These assumptions will influence how evaluators will need to prepare their estimate of magnitude. The principal assumptions in the decision analysis models are:

1) That TWA is sufficiently correlated with a bioactive metric so that it is an adequate exposure metric to use for the assessment.

2) There are four threshold assumptions:

a) There is no threshold for the relationship between exposure to EMFs and risk of adverse health effects. Another way of saying this is that there is no level of exposure to EMFs that does not increase the risk of disease, at least to some extent.

b) "Linear effects" (glossary) begin at 2 mG, 5 mG, or 10 mG.

c) That the relationship between exposure to EMFs and adverse health effects (if real) would be likely to have a plateau (the level at which risk no longer increases). This assumption is necessary to prevent individual risks of common diseases from exceeding 100%. This assumption means that, above a certain level, the risk of adverse health effects does not increase with further increases in TWA.

d) The exposure metric to be used, TWA, has averaging times ranging from a few hours to 24 hours. There are other metrics in the decision model, but there is insufficient evidence to develop dose-response curves for them.

The models require, as one of the input data, the slope of the linear part of the dose-response function. To make this more intuitive, instead of requesting the usual relative risk increment for unit measure of exposure, the model asks the user to input his or her best estimate of the ratio of the risk of a subject exposed to 2 mG compared to that of the risk of a subject totally unexposed. This is approximately equal to the dichotomous odds ratio (OR) reported in the several epidemiological studies using a cut point of 2 mG. However, for some values of the relative risk at the 2 mG level, the risk ratio and the dichotomous OR at 2 mG are significantly different. Moreover, not all studies report a dichotomous odds ratio at 2 mG. To obviate these problems, EMF program staff have used computer modeling to produce a series of tables relating these two measures of risks for different environmental exposure distributions.

To address the decision analysis model, DHS evaluators will estimate the slope of the exposure-response curve and assess evidence, if any, for or against the existence of a threshold and for or against the existence of a plateau. Evaluators should define what the plateau would be, if there is evidence for it. Evaluators are to comment on whether they have identified any empirical basis to use another model for the relationship between exposure and disease response. Evaluators should also discuss assumptions for the shape of a dose-response curve for TWA with thresholds. Evaluators should also discuss the likelihood that TWA is a poor surrogate for some other attribute of the EMF mixture and what the practical consequences of that assumption would be. The evaluator should comment on evidence, if any, that measurements were not taken at the vulnerable receptor organ, or that EMFs only work during circadian or developmental windows of vulnerability. Since it has been argued that any effect of EMF would vary as the square of the field, the DHS evaluators will address the merit of this argument and its compatibility with the evidence.

This gives five risk functions to evaluate:

for linear models:

a) no threshold

b) 2 mG threshold

c) 5 mG threshold

d) 1o mG threshold

a model based on the square of the TWA

There should also be comment about especially vulnerable subgroups if any. Any model must meet these constraints.

- since all of the diseases to be evaluated existed before the widespread introduction of 50 to 60 Hz electricity there must be a residual risk even when exposure to EMF is zero
- when applied to environments in which epidemiological studies have been conducted, such as Denver or Los Angeles, the model must yield results consistent with the results reported in these studies
- the model cannot lead to any individual having a probability of disease greater than 100%, no matter what the exposure
- when applied to highly exposed populations, such as some occupationally exposed workers, the model cannot predict a rate of disease greater than that observed

Attributable Population Burden

After estimating the magnitude of the association between exposure and response, the next step is to apply these results to the population of California, using the best available estimates of exposure, to estimate the burden of disease that may be associated with EMF exposure in the population.

Ideally, we would identify the biologically active attribute(s) of the EMF mixture, determine appropriate units of measure, and establish a precise relationship between this and the surrogate metric used in the epidemiological studies. This would then allow us to focus on exposure to the appropriate bioactive agent. We would then need to establish the dose (how much of the exposure was actually absorbed by the subject) and the dosing schedule (how this dose was distributed in time and the relationship of this time distribution to the time distribution capable of effecting adverse biological changes). The information available to evaluators is likely to be less detailed than would be ideal. We have one exposure metric, time-weighted average field strength (TWA).

The theoretical attributable population burden is derived by applying a dose-response curve to the number of persons in each exposure category to determine the amount of disease expected to result from exposure above that expected if the entire population had been in the lowest exposure bin. The resulting number represents the annual number of cases that could be avoided if we were certain that the epidemiological association was causal, the shape and slope of the dose response curve exactly right and the exposure removed from the population. For our purposes, the best available estimate of current personal exposures comes from Zaffanella's recent 1000-person study in the United States which provides personal 24-hour monitoring data and the proportion of the population which can be found at various levels of TWA.[15]

Evaluators should report attributable population burden for the relevant disease outcomes both with certainty weighting and without. We will calculate these for all five risk assumptions (describing the proportion of theoretical cases that are generated for 0 to1 mG, 1 to 2 mG, 2 to 3 mg and 3 mG and above) so that decision-makers can see the consequences of uncertainty about the shape of theoretical dose-response curves. Implicitly, these calculations will also convey information on the population impact of intervention, as they will provide estimates of how many cases would be avoided if all exposure about 1 or 2 mG were eliminated. These estimates will be presented with and without weighting by the degree of confidence.

Lifetime Attributable Risk at 90th Percentile Exposure

Some regulatory decisions and voluntary individual decisions are influenced by the risk accumulated from a lifetime of exposure. In California, Proposition 65 labeling is triggered if the accumulated theoretical risk from a lifetime of exposure exceeds 1 per 100,000.

If one has the range of relative risks conveyed by the 90th percentile of exposure, one can apply these to the schedule of age-specific baseline rates of the disease in question to estimate the probability of escaping that disease in each year of life, with a 90th percentile exposure or with zero exposure. One can then calculate the probability of escaping that disease in a lifetime with the two exposure scenarios, and then calculate the complement, the theoretical probability of getting this disease with a 90th percentile exposure versus zero exposure. The difference represents the added lifetime risk.

The DHS evaluators will estimate the lifetime attributable risk for populations in California using the five risk function assumptions. We will present these estimates weighted by the degree of confidence as well as without this weighting.

VII. Summary of Potential Risks

The population attributable burden and lifetime attributable risks provide our best estimate of the overall theoretical burden of ill health to the population and to individuals highly exposed throughout their lifetime. Provide a summary and explanation of these.

Provide a table that shows the estimates relevant for the various diseases considered in the assessment. The following table may be an appropriate model.

Table 3. Disease-by-disease estimates of theoretical risks from EMFs

Disease	Magnitude of Association	Attributable Population Burden	Individual Lifetime Risk	Confidence-Weighted Attributable Population Burden
	slope of line for TWA and rate ratio	estimate of annual number of cases above background	estimate of individual risk from 90th percentile	estimate of number of cases adjusted by degree of confidence

In our decision models, these estimates would be multiplied by the degree of confidence. Thus, if barely detectable epidemiological results suggested the possibility of 100,000 deaths from EMFs, but we had only 1% confidence in this, our confidence-weighted attributable population burden would be 1% x 100,000, which equals 1000. Because we recognize that such population weighting has ethical implications and may not be appropriate in many contexts and because we also want our results to be comparable to those developed by US EPA and Cal EPA, DHS evaluators will present both confidence-weighted and -unweighted estimates for these terms

We recognize that certainty weighting combines very different types of measures and cannot be used uncritically. We intend to use these estimates to explore policy options in the decision models but not to advocate this approach in other contexts.

V. Risk Communication Statement for Each Disease or Condition

Since risk evaluations can be framed to produce different responses, DHS plans to provide a recommended summary statement that best captures the "bottom line." This summary statement will be framed to fit both regulatory and individual decision-makers. It should explicitly warn against selective out-of-context quotations from other parts of the document, particularly from the pro and con arguments preceding the explicated final judgment in each disease section. It will be designed to avoid inducing either inappropriate complacency or over-reaction.

VIII. Appendix on Mitigation

To estimate accurately the population burden and the effectiveness of exposure mitigation, we would need to know how the distribution in the population of the true exposure (as opposed to that of the surrogate metric) and the distribution of exposure events over time. For example, if exposure were bioactive only if received during sleep, with an intensity never dropping below 2 mG for a period of at least five minutes, we would find that the frequency distribution of such events is probably correlated to, but substantially different from, the distribution of point-in-time (spot) measurements used routinely in epidemiological studies.

The decision analysis focuses ultimately on the expected number of cases of various diseases before and after mitigation has changed the distribution of exposure in the population.

For decision analysis, this attributable population burden can be multiplied by our degree of certainty that the association was causal. This is used to determine the benefits of mitigation. A good decision analysis tries to estimate what the exposure distribution would be with the mitigation options being evaluated.

If we had this information, we could calculate how this distribution is changed by a given mitigation strategy and, finally, how the population burden would be reduced by mitigation.

Even if one was convinced that the associations between disease and occupying certain job categories were causally due to EMFs one might have residual uncertainties about what attribute or dosing schedule of EMFs ought to be modified in that job. Is it the 60 Hz attribute that is of interest or the transients? Should we lower the 24-hour average exposure or do we need to avoid even brief high exposures? These are the issues that should be dealt within this section.

Organize the discussion in terms of broad categories of mitigation such as: increasing distances from power lines, burying power lines, measures which result in lowered time-weighted averages, measures which result in trading prolonged moderate exposures for brief high exposures (by placing necessary sources in infrequently used locations).

For each mitigation class, review the biophysics, mechanistic studies, whole-animal studies and epidemiology for different diseases that might relate to your degree of certainty that this class of mitigation options might be effective. Pay particular attention to attributes such as transients, or dosing schedules such as short high exposures that might not be affected equally by all mitigation classes. Deal explicitly with the likelihood that a mitigation class would move people into or out of a bioactive window.

ABBREVIATIONS

CalEPA	California Environmental Protection Agency
CPUC	California Public Utilities Commission
DHS	Department of Health Services
IARC	International Agency for Research on Cancer
PHI	Public Health Institute
RFP	Request for Proposals
SAC	Stakeholder Advisory Consultants
SAP	Science Advisory Panel
US EPA	US Environmental Protection Agency
WHO	World Health Organization

GLOSSARY

This document is pivotal to the conclusion of the EMF program, a process blending risk evaluation, exposure assessment and decision analysis. In order for these elements to come together, it is essential that key terms be used with the same meaning. This is not necessarily an obvious fact. Technical terms may have correct, even strict, yet differing definitions in different disciplines. For example, epidemiologists Carmines and Zeller proposed the following definition for "measurement": linking abstract concepts to empirical indicators.[20] This would horrify a physicist, who learns from the first day of his/her training that measurement is defined as a simple arithmetic operation, the ratio between two homogenous quantities, one of which is chosen as the measurement *unit*. For these reasons, we expect that different readers may disagree with the definitions given below. However, we believe that these are the best definitions for the limited scope explained above, and we will use them in this document, the risk evaluation itself, the policy analyses and the policy integration projects.

Some of these definitions were adapted from a recent report from the National Institute of Environmental Health Sciences.[1]

Bayes theorem – (applied to EMF) the updated or modified odds that EMF causes disease after seeing new evidence are equal to the odds that EMF causes disease before seeing the evidence, multiplied by the relative likelihood of the pattern of new evidence.

Bayesian network – a quantitative method for developing an estimate of the certainty of a relationship, using evidence from multiple sources. After setting a series of *a priori* relative likelihood for streams of evidence and their inter correlations, a complicated computer algorithm updates all the relative likelihoods in the network as actual results are entered. The relative likelihoods in this scheme depend both on the strength and quality of the evidence.

attributes – detailed physical properties of electric or magnetic fields, such as the intensity, frequency spectrum, or polarization.

confidence interval – see "statistical significance" *decision analysis* – a framework used to systematically analyze the impact of different conditions and assumptions, in order to make difficult decisions. Decision analysis typically constructs a model to describe the various aspects of the decision and uses sensitivity analysis to determine which aspects are important.

dose – a toxicological term for the amount of a chemical or physical agent delivered to the body or to an organ in the body at a point in time or over a defined period of time. For EMFs, the concept of an exposure metric is usually used instead of dose because the target organs and the mechanism of delivery of the dose are not well understood.

dosing schedule – the way that exposure is delivered over a period of time, usually a day, week, or month.

ELF (extremely low frequency fields) - EMF with frequency ranges from 3 to 3000 Hz.

EMF or EMFs (electric and magnetic fields) - the combination of electric and magnetic fields in the environment.

electric distribution lines – lines that carry electric power from the power grid to neighborhoods and business areas.

environmental EMFs – the types of fields that people might expected to be exposed to in residential, school, or business environments.

environmental levels (of EMFs) - time-weighted average (see below) values exhibiting a strongly skewed distribution, with median values around 1 mG in residential environments and 1.5-2 mG in most occupational environments, but with 95^{th} percentile values (several milliGauss) in residential environments and tens of milliGauss in some of the most exposed occupations.

exposure – the amount of a chemical or physical agent (EMFs, here) in the environment that a person comes into contact with over some period of time.

exposure metric – a single number that summarizes an electric and/or magnetic field exposure over a period of time. An exposure metric is usually determined by a combination of the instrument's signal processing and the data analysis performed after the measurement.

extremely low frequency fields (ELF) – frequency range from zero to 300 Hz

false positive rate (FP) – the probability that a test will yield a positive result when in reality the accurate result should have been negative.

Independent System Operators – an organization created to manage the transmission power grid in the state of California. After the deregulation of the electric utilities in 1997 this organization is to carry out some functions previously carried out by electric utilities.

intermittent fields – fields whose *rms* (below) vector magnitude changes rapidly, with a time scale of seconds. In contrast to transients, intermittent fields may have high levels for longer times and are generally in the ELF frequency range.

human evidence – information about the relationship between exposures to agents and disease or physical changes that comes from studies on people. Such studies often occur in workplaces, where workers may be exposed to higher concentrations of hazardous agents than the general population. Some human evidence is experimental.

Hertz (Hz) –cycles per second

kiloHertz (kHz) – one thousand cycles per second (1000 Hz)

linear model (with no threshold) – a linear relationship is one where the amount of adverse effect increases whenever the amount of the agent increases. In our decision models we assume that the relative risk increases linearly with time-weighted average (TWA, below). In such cases, for any increase in the exposure to the agent, there would be an increase in disease. If there is no threshold, this means that there is no exposure to the agent that is completely safe and that does not increase the risk of disease to some extent.

low frequencies (LF) – frequency range from 30 to 300 kHz..

metric – see "exposure metric"

milliGauss (mG) – a measure of the strength of a magnetic field. A typical living room would be measured at 0.5 mG.

mitigation – steps taken to reduce exposure to EMFs or attributes of the EMF mixture. Examples of mitigation might include placing power distribution lines underground or changing the configuration of wiring in homes to reduce field production.

one-tailed test – a test of statistical significance used when only one side of the alternative to the null hypothesis is being considered. For example, if the null hypothesis is that EMF does not have health effects, the alternative hypothesis, EMF does have health effects, has two sides: EMF has beneficial health effects or EMF has adverse health effects. If we are only considering the possibility of *adverse* health effects, we should use one-tailed significance tests.

power frequency – the frequency at which AC electricity is generated. For electric utilities, the power frequency is 60 Hz in North America. Electric power is 50 Hz in much of the rest of the world.

probability elicitation – a process of drawing from participants their best estimates of the likelihood that some condition is true.

rms (root mean square) – square root of the average of the squared instantaneous intensities during one cycle of an alternating current.

static field – a field whose direction or intensity does not vary with time.

statistical significance – a measure of the probability that a certain observation is *not* due to chance. Statistical significance can be expressed by a "p-value" indicating the probability that a result different from the null is due to chance. For example, to say that a study indicates that two factors are correlated, with a "p-value" of 0.01, means that there is only a 1% probability that the two quantities are in fact *not* correlated. A confidence interval is another way of measuring statistical significance which also places bounds on the extent that chance may alter a result. For example an estimated value of 3 with a 95% confidence interval of 1.1 - 8.2 means that we are 95% confident that the true value is no less than 1.1 and no more than 8.2.

surrogate (exposure surrogate) – an easily accessible way of measuring that is a good substitute for a complicated way.

time-weighted average (TWA) – a weighted average of exposure measurements taken over a period of time with the weighting factor equal to the time spent in the place measured.

transients – brief, microsecond bursts of high frequency fields, usually resulting from mechanical switching of AC electricity. Much shorter than intermittent fields.

transmission lines – power lines that carry large quantities of power large distances. These are high voltage lines that comprise the state transmission power grid.

true positive rate (TP) – the probability that a test will yield a positive result when a positive result is the accurate one.

two-tailed test – a test of statistical significance used when both sides of the alternative to the null hypothesis are being considered. For example, if the null hypothesis is that EMF does not have health effects, the alternative hypothesis, EMF does have health effects, has two sides: EMF has beneficial health effects or EMF has adverse health effects. If we are want to determine whether *any* health effects exist, irrespective of whether they are beneficial or harmful, we should use two-tailed significance tests.

ultra low frequency (ULF) – the frequency range below 3 Hz.

very low frequency (VLF) - the frequency range from 3 to 30 Hz.

wire codes – a way to classify configuration of power or distribution lines to estimate potential for exposure to EMFs. In this assessment, "wire-coding" will be regarded as a surrogate for time-weighted average fields, rather than as an exposure metric in its own right.

APPENDIX ONE Science Advisory Panel (SAP) Members

The area of expertise of Science Advisory Panel members is given, along with his or her affiliation.

physics Stephen L. Brown, PhD, Director, R2C2 Risks of Radiation and Chemical Compounds
4700 Grass Valley Rd., Oakland, CA 94605

ethics/law Carl Cranor, PhD, Associate Dean, College of Humanities, Arts and Social Sciences
University of California Riverside
Deans' Office, Humanities & Social Sciences Bldg.
NE Wing, 3rd Fl, Riverside, CA 92521

molecular biology James E. Cleaver, PhD, Program Director, Cutaneous Oncology
University of California San Francisco, San Francisco, CA 94143

policy analysis Shan Cretin, PhD, MPH, Senior Analyst
RAND Health Program
P.O. Box 2138, Santa Monica, CA 90407-2138

toxicology Michael S. Denison, PhD, Professor of Environmental Toxicology
University of California, Davis
4241 Meyer Hall, Davis, CA 95616

epidemiology Hal Morgenstern, PhD, Professor of Epidemiology
UCLA School of Public Health
Los Angeles, CA 90095-1772

biostatistics Charles P. Quesenberry, PhD, Senior Biostatistician
Kaiser Foundation Research Institute, Division of Research
3505 Broadway, Oakland, CA 94611-5714

exposure assessment Robert Spear, PhD
University of California Berkeley School of Public Health
Berkeley, CA 94720

Dr. Spear resigned from the SAP in 1999 and has been replaced by Dr. McKone:
exposure assessment Thomas E. McKone, PhD, Adj Professor
Lawrence Berkeley Laboratory, University of California Berkeley School of Public Health
140 Warren Hall, Berkeley, CA 94720

cell biology James David Tucker, PhD, Senior Staff Scientist, Biology & Biotechnology Research Program
Lawrence Livermore National Laboratory
University of California Berkeley
P.O. Box 808, L-452, Livermore, CA 94551

oncology Jan Van Tornout, MD, MS, Assistant Professor, Pediatrics & Preventive Medicine
Children's Hospital LA-USC
MS #54, 4650 Sunset Blvd, Los Angeles, CA 90027

chair Warren Winkelstein, MD, MPH
University of California Berkeley School of Public Health
Berkeley, CA 94720-7360

APPENDIX TWO

How to Express Quantitively a Change in Confidence Brought about by Reviewing Evidence

The guidelines incorporate elements of a "Bayesian" approach to review of scientific evidence.[5] The basic elements of this approach are to define *a priori* one's degree of confidence that an association between exposure to an agent and a disease outcome is truly causal in nature. Then one uses a quantitative treatment of available evidence to adjust this degree of confidence either up or down. The guidelines propose a qualitative analogue of this process which will be more transparent to the scientific community and the general public than a fully quantitative treatment would be. However, this appendix explains how the quantitative process would be done so that the quantitative rationale for our qualitative approach is made explicit.

The hypotheses we wish to contrast are:

1) The ranges of usual environmental and or occupational exposures are contributory causes and partially explain the epidemiological associations with certain diseases. The exact exposure metric is not known, but if causal would be correlated to the time-weighted average (TWA) of the root mean square (rms) of the magnetic flux density or of the electric field strength.

2) They are not contributory causes of these diseases and do not explain the epidemiological associations seen.

Probablistic (Bayesian) causal inference views the reasoning process as follows: On the basis of general knowledge one starts out with an initial or "prior" degree of confidence, which one can express as the prior or "initial odds." The "odds" are defined as: (1) the probability that EMFs cause disease divided by (2) the probability that EMFs don't cause disease. If the first term were 80% and the second term 20%, then the odds would be 80 to 20 or 4 to 1.

One then conducts relevant studies and looks at the pattern of evidence. One contrasts the likelihood of this observed pattern of evidence if EMFs did cause disease to the likelihood of this pattern if EMFs did not cause disease. One takes the ratio of the two likelihoods to get a "relative likelihood."

By multiplying the relative likelihood by the prior odds one derives "updated" or "posterior odds." If the pattern of evidence is much more likely if EMFs cause disease, then the relative likelihood will be a big number and the posterior odds will be much bigger than the prior odds. If the pattern of evidence is the kind one would see if EMFs didn't cause disease, the relative likelihood will be a fractional number less than 1 and the posterior odds will be smaller than the prior odds.

The terms involved in this procedure can be written as:

$$\text{prior odds:} \quad \frac{P\ (\text{cause})}{P\ (\text{not cause})}$$

$$\text{relative likelihood of this pattern of evidence:} \quad \frac{P\ (\text{this evidence} \mid \text{cause})}{P\ (\text{this evidence} \mid \text{not cause})}$$

$$\text{posterior odds:} \quad \frac{P\ (\text{cause} \mid \text{this evidence})}{P\ (\text{not cause} \mid \text{this evidence})}$$

The relative likelihood would be read out aloud as: "The probability of this evidence, given the hypothesis that EMFs cause disease, divided by the probability of this pattern of evidence, given the hypothesis that EMFs don't cause disease."

Bayes Theorem can then be written as:

$$\frac{P(\text{cause} \mid \text{this evidence})}{P(\text{not cause} \mid \text{this evidence})} = \frac{P(\text{this evidence} \mid \text{cause})}{P(\text{this evidence} \mid \text{not cause})} \cdot \frac{P(\text{cause})}{P(\text{not cause})}$$

This is the relationship:

> If the middle term, the "relative likelihood," is 1, the evidence does not change the odds. If the relative likelihood is bigger than 1 it modifies the odds upward. If it is less than 1 it modifies the odds downward. Of course, all three terms are conditional on the background body of relevant scientific knowledge available before starting to consider this problem.

For clinical tests the relative likelihood can be computed using the known sensitivity and specificity of the test. For example, if we knew that a test correctly identified 12% of human cases of a disease and falsely identified 6%, the relative likelihood conveyed by a positive test would be 12 divided by 6 = 2. So a positive test would increase the odds by a factor of 2. Conversely, a negative test would convey a relative likelihood of (100 minus 12) divided by (100 minus 6) = 88 divided by 94, or 0.94.

The terminology for laboratory tests can be carried over to bodies of evidence and related to the Bayes Theorem terminology and to terms used in hypothesis testing, as in Table A1 (glossary).

Table A1. Equivalent terms from Bayes Theorem, hypothesis tests and laboratory tests

	THE TRUE SITUATION	
	alternative hypothesis: EMFs cause disease	**null hypothesis: EMFs don't cause disease**
positive evidence	true positive number (TP) true positve rate (sensitivity) = (TP) / (TP+FN) = P (Pos Ev | cause)	false positive number (FP) false positive rate (Type I Error Rate) = (FP) / (FP+TN) = P (Pos Ev | not cause)
negative evidence	false negative number (FN) false negative rate (Type II Error Rate) = (FN) / (TP+FN) = P (Neg Ev | cause)	true negative number (TN) true negative rate (specificity) = (TN) / (FP+TN) = P (Neg Ev | not cause)

The relative likelihood conveyed by positive evidence $= \frac{\text{TP rate}}{\text{FP rate}}$ or $\frac{\text{sensitivity}}{\text{Type I error rate}}$

The relative likelihood conveyed by negative evidence $= \frac{\text{FN Rate}}{\text{TN Rate}}$ or $\frac{\text{Type II error rate}}{\text{specificity}}$

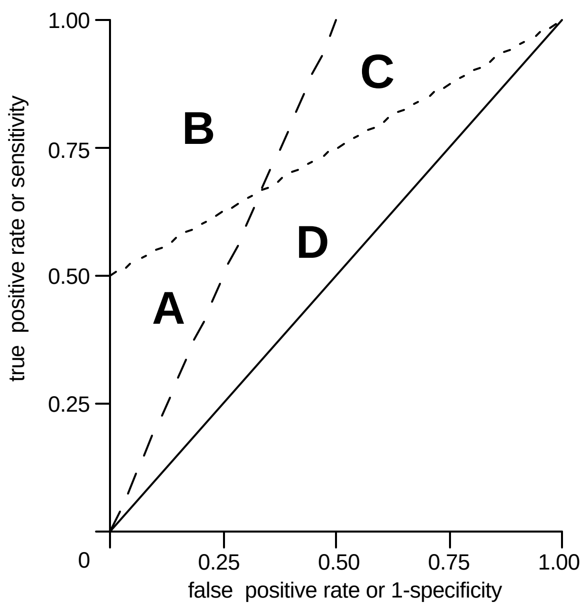
Table A2 (below) provides examples of "strengthening or weakening," "predominantly strengthening," "predominantly weakening" or "uninformative" evidence.

Table A2. Examples of how sensitivity and specificity of tests affect interpretation of results

Evidence Type	True Positives	False Positives	Relative Likelihood from a Positive	Relative Likelihood from a Negative
strengthening or weakening	0.88	0.02	88/2 = 44	12/98 = 0.12
predominantly strengthening	0.12	0.06	12/6 = 2	88/94 = 0.94
predominantly weakening	0.97	0.88	97/88 = 1.1	3/12 = 0.25
uninformative	0.94	0.94	94/94 = 1	6/6 = 1

Every possible combination of false positive (FP) and true positive (TP) has two relative likelihoods. The first is that conveyed by a positive result for the test with that TP/FP combination. The second is that conveyed by a negative result from that false negative (FN) / true negative (TN) combination. The particular examples of evidence types in Table 1 can be visualized if we create a graph with the frequency of true positives or sensitivity on the vertical axis and the frequency of false positives or 1-specificity on the horizontal axis (below).

Figure A1 Plot of true positives and false positives and the four classes of evidence



For illustrative purposes we will require the likelihood ratio conveyed by a positive test to double our odds or that conveyed by a negative test to halve them before we consider either to be informative. Other criteria could be justified by a particular decision context, with particular penalties for false positives and negatives, but the same general insights mentioned below would be derived.

The zones of TP/FP spaces, which delineate families of tests with these properties, are formed after laying down two lines. The first line indicates those combinations of TP and FP in which a positive test conveys a likelihood ratio greater than or equal to 2. The second line is likelihood ratios conveyed by a negative test which are less than or equal to one half.

It is intuitively obvious that the formula for the first line is TP = 2 FP, represented by the long-dashed line. It can be shown that the formula for the second line is TP = 0.5 FP + 0.5, represented by the fine-dashed line. Their intersection occurs at TP = 0.666 and FP = 0.333 and generates four areas in the TP/FP space.

Zone A represents the family of tests or body of evidence for which a positive result conveys a likelihood ratio which at least doubles our odds, but that of a negative result would not cut our odds in half. The evidence in Zone A is *predominantly strengthening*. Zone B is the family of tests for which positive results will at least double our confidence and negative evidence will at least cut it in half. This can thus both *strengthen or weaken our confidence* substantially. Zone C is the body of evidence for which a negative test result will cut our odds in half, but a positive result will not double our odds. This family of evidence can *only weaken* our confidence. Zone D is that family of tests or body of evidence whose combination of false positives and true positives convey relative likelihoods weaker than the above mentioned criteria, the zone of *uninformative* tests.

The concepts described above were originally applied to the fields of laboratory tests and medical diagnosis. For example, a patient might present to the emergency room with abdominal pain. The physician has initial odds that the pain is caused by one of several things, some of which require surgery and a few which do not. A laboratory test with a given set of false positives and negatives conveys one likelihood ratio if positive and another if negative. These likelihood ratios can modify the doctor's odds and help her select a surgical or non-surgical intervention.

How could these ideas be applied to epidemiological studies or animal bioassays instead of laboratory tests? Here the risk evaluator has initial odds that a particular agent causes disease X. Depending on the expected strength of association and the study design, one could describe as "large," "medium" or "small" the expected probability of a positive result if the agent really does cause the disease, and the expected probability of a positive result despite the fact that the agent doesn't really cause the disease. On this basis, one can reason about whether this kind of study will provide strengthening *and* weakening evidence or whether it fits in one of the other three families of evidence mentioned above. Examples have been given in the main body of the guidelines.

The discussion above could be applied to individual studies or to a whole stream of evidence. If these studies are independent and their results uncorrelated, it can be shown (2) that, given patterns of evidence E1 and E2 from two studies, that:

$$\frac{P\,(E1 \cap E2 \mid \text{cause})}{P\,(E1 \cap E2 \mid \text{not cause})} = \frac{P\,(E1 \mid \text{cause})}{P\,(E1 \mid \text{not cause})} \cdot \frac{P\,(E2 \mid \text{cause})}{P\,(E2 \mid \text{not cause})}$$

That is the relative likelihood conveyed by the combined results is the same as the product of the relative likelihoods from each study alone. The same point could be argued for relative likelihoods conveyed by independent streams of evidence. We are not proposing to discuss the relative likelihoods conveyed by each study within a stream of evidence since to do this properly would require a complexity which would make our reasoning process difficult for most people to follow. We will discuss the relative likelihood conveyed by the mechanistic, whole animal and human evidence and keep in mind as a heuristic that if they were completely independent, these relative likelihoods would be the weights assigned to these streams of evidence and that they would be multiplied one by the other and by the prior odds.

The intercorrelation of one stream of evidence to another could be treated quantitavely through the use of a Bayesian “network.”[21] When we discussed this full-blown, highly quantitative approach with our stakeholders and other experienced risk assessors, two problems were pointed out. First, scientists are not likely to agree on the numerical values for all the needed parameters. Second, even if this were possible, only those with extensive training in quantitative methods would be able to evaluate the process by which the final degree of confidence was produced. It would also be more difficult to double check for stakeholders with less resources.

Hutchison and Lane[22] have discussed using a Bayesian approach to determining if an untoward drug side reaction was due to the drug or some other cause. They recommend using a predetermined set of questions about the evidence (“explicitness”); not ruling out any evidence from consideration (“completeness”); and considering the likelihood of the evidence if there were a hazard and the likelihood of the evidence if there were no hazard (“etiological balancing”). They recommend a method for moving from the pattern of evidence to the degree of confidence of causality which can be understood (“transparency”). We strive through our pro and con and summary arguments to achieve transparency and through a thorough characterization of the evidence to achieve the other desired characteristics of a causal evaluation.

APPENDIX THREE

CRITERIA FOR EMF HEALTH RISK ASSESSMENT

MH Repacholi[1] and E Cardis[2]

[1]World Health Organization, Geneva, Switzerland

[2]International Agency for Research on Cancer, Lyon, France

Published in: Radiation Protection Dosimetry 72: 305-312, 1997

ABSTRACT

The International EMF Project was established at WHO in 1996 to provide a forum for a coordinated international response to health issues raised by exposure to electric, magnetic and electromagnetic fields (EMF). Research on EMF has been ad hoc and in many cases uncoordinated. Unreplicated research has been placed at the same level as high quality research that establishes results in a scientifically valid manner. Because of this the EMF issues have now reached a high level of concern among the general public and workers. This needs to be addressed at the international level, since the problem is truly global in nature. Research objectives are needed with a clear focus to improve our database of science used for health risk assessments. This paper indicates how the International EMF Project will evaluate scientific reports, identify the scientific database needed to make health risk assessments, and assess health hazards using established IARC criteria.

Running title: EMF Criteria

Author to contact for editorial changes:
Dr Michael H Repacholi
Responsible Officer,
Radiation Protection and Global Hazards Assessment
Office of Global and Integrated Environmental Health
World Health Organization, CH-1211 Geneva 27, Switzerland
Tel: +41 22 791 3427, Fax:+41 22 791 4123, E-mail: repacholim@who.ch

Published: Repacholi MH, Cardis E (1997): Criteria for EMF Health Risk Assessment. Radiat Prot Dosim 72: 305-312.

INTRODUCTION

Biological effects and possible health consequences of exposure to electromagnetic fields (EMF) need to be assessed according to an appropriate set of guidelines. Through the International EMF Project(1), WHO is collaborating with its specialised agency on cancer research, the International Agency for Research on Cancer (IARC), and other international organizations, including the International Commission on Non-Ionizing Radiation Protection (ICNIRP), governmental agencies and independent research institutions, to assess health effects of exposure to static and time varying electric and magnetic fields in the frequency range 0 - 300 GHz. The Project incorporates a framework for identifying gaps in knowledge, establishing a research agenda to enlarge the scientific database and completing reviews of the literature in a manner that leads to scientifically defensible conclusions on possible health risks from EMF exposure. The International EMF Project provides a global focus on the EMF issues and facilitates progress towards scientifically acceptable solutions. It is particularly important that the scientific community, general public and workers are reassured that the Project is addressing all the health concerns in a logical and coordinated manner so they will have confidence in the final results.

One of the greatest problems in assessing health risk has been the lack of consistency of results in the EMF scientific database. Results of many studies have not been replicated and so reports which could have important implications for health have remained unsubstantiated. While exact replication of studies may not be necessary, additional studies are needed to support the same conclusions. A major goal of the International EMF Project will be the identification of a research agenda, the results of which would provide a better scientific database on which health risk assessments can be made, and encouragement of funding agencies to support this research. The results of research from this agenda will be added to reviews of published literature prior to publication. Major independent reviews of the literature will assist in this process.

Another objective of the International EMF project is to evaluate health risk from EMF exposure. This paper provides information on how these evaluations will be carried out and particularly the criteria on research needs, and the evaluation of scientific reports and health hazards from EMF exposure.

SCIENTIFIC DATABASE NEEDED TO EVALUATE HEALTH RISK

The database needed to evaluate whether exposure to any physical or chemical agent produces a carcinogenic risk has been described by the International Agency for Research on Cancer(2) and has been elaborated by Cardis and Rice(3). Effectively the same type of scientific database can be used for determining any risk to health from EMF exposure. The following describes the database for EMF which will be used in the International EMF Project. Studies reporting both positive and negative effects will be critically evaluated to determine whether the effect studied is related to EMF exposure. Criteria for this evaluation are described below.

Studies in Humans

Epidemiological studies contributing to the evaluation of EMF health effects are of two main types: Cohort studies and case-control studies. While there are other categories such as correlation studies, randomised clinical trials and case reports in humans, they are rarely available for EMF effects nor do they have sufficient power to be useful in health risk evaluation. Cohort studies relate estimates of individual EMF exposures to the occurrence of the studied health effect(s) in a group of individuals and provide an estimate of relative risk (ratio of incidence or mortality in those exposed to the incidence or mortality in those not exposed) as the main measure of the association. Case-control studies compare the exposure of individuals with and without the disease.

Exposure Assessment

A major concern with EMF epidemiological studies has been exposure assessment. Since laboratory studies have been unable to establish mechanisms for health effects occurring at low or "environmental" EMF exposure levels, or any clear concept of the dose metric at these levels, exposure assessment has been determined using various methods. In many cases, surrogate or proxy measures have been used as an index of EMF exposure. Examples of these measures that have been used for low frequency (50/60 Hz) fields are given below.

Magnetic field measurement: Spot (a single measurement in a given position), peak (maximum field) and 24-hour average (placing a magnetic field measuring device in a room for 24 hours and taking the time-weighted average of the reading) field measurements have been performed in residences in some of the major studies as estimates of personal exposure. This method may take some account of fields from house wiring and domestic electrical appliances, but not of exposures received away from residences.

Distance to power lines: Proximity of residences to high voltage power line corridors has been used as a measure of a person's magnetic field exposure. This exposure metric assumes that high voltage transmission lines are the dominant contributor to exposure and so do not account for field contributions from within or away from residences.

Wire Codes: The original study conducted by Wertheimer and Leeper(4) used a combination of a number of factors that related to the amount of electrical current flowing through wires or conductors. Since the magnitude of the current relates to the strength of the magnetic field, the type of wiring (distribution or transmission line, number and thickness of wires) and distance of the wiring from the residence was used as a surrogate for the measure of electric and magnetic field exposure. This technique is called "wire coding", and, in a more refined form, has now been used in a number of subsequent studies(5). This method has the advantage of being able to classify a home as high, medium or low current configuration from the exterior. However, it cannot account for domestic field exposures unless additional measurements are taken.

Historic magnetic fields: Recent studies, eg Feychting and Ahlbom(6), have used power company records and maps to calculate the magnetic field strengths that would have been produced in the past from high voltage transmission lines. These fields are calculated using historical line current loadings, configuration of the conductors, and distance of the residence from the line. Typically, historical measures of field exposure are determined at the time of diagnosis of the cancer or as the average magnetic field for a number of years prior to diagnosis. When this method is checked against measured magnetic fields at a given location, they correlate reasonably well. However, this technique cannot account for a person's magnetic field exposure from local distribution lines (even though they may be underground), or determine the contribution from household wiring and appliances. Further, there is no way of checking the accuracy of calculated historic fields.

Job classification: Many occupational studies have used various combinations of job title, type and duration of work, and workplace field levels to categorise exposure or compile an exposure index. This method assumes that occupational exposure far exceeds residential or other non-occupational exposures, and so no account of these are normally taken.

For epidemiological studies involving radiofrequency field (RF) exposure, similar surrogates or direct measures have been used. They vary from job titles with some local field measurements to distance from RF sources. Some studies have attempted to estimate the specific absorption rate (SAR) for the study populations. It is generally agreed that RF exposure in certain occupations far exceeds those in residences. The exception would be during use of such devices as mobile telephones. Here near field RF exposures exceed any environmental levels.

In order for the evidence from studies to be evaluated, the method of exposure assessment should be reported in detail. If a surrogate is used, it needs to be documented and validated. Details of exposure metrics should be provided and preferably address issues such as the field strengths, how they were measured, their characteristics, how or if transients were considered, night-time versus daytime exposure, or domestic (including non-occupational exposures: shopping, schools) versus occupational exposure. This is extremely important when accumulating evidence for causality. A good description of wire codes and their relationship to measured and historic magnetic fields, and prediction of field exposure classification or personal exposure, is given in NRC(6). Further information on RF field dosimetry in epidemiological studies is given in Repacholi(7).

Study Quality

When evaluating the quality of human studies, it is not necessary to assess in detail all reports. Those judged to be inadequate or irrelevant to the evaluation are generally omitted. Brief mention may occur when the information is useful to supplement other reports or when they provide the only data available.

It is necessary to take into account the possible roles of bias, confounding and chance in the interpretation of study results. Bias is the operation of factors in the study design or execution that lead erroneously to a stronger or weaker association than exists between exposure and the disease under study. Confounding occurs in situations where the relationship with the disease is made to appear stronger or weaker than it truly is as a result of an association between the apparent causal factor and another factor that is associated with either an increase or decrease in the incidence of the disease. Lack of clarity in the reporting of these factors can decrease the credibility and final weight given to the results of the study.

For epidemiological studies to be informative for the evaluation of health risks related to EMF exposure the following aspects should be addressed:

1. Hypotheses to be tested, study population, disease(s) and exposure assessment should be well defined at the outset by researchers. Cases of disease should be identified in such a way that it is independent of EMF exposure, and exposure should be assessed in a way that is not related to disease status.

2. Researchers should take into account, in both the study design and analysis, any variables (confounders) that could influence the risk of the disease and may also be related to EMF exposure. While there are few known confounders for EMF study diseases of interest, these should be dealt within the study design, such as by carefully matching cases and controls, and in the analysis by statistical adjustment.

3. In EMF studies, categorizing the study population into different levels of exposure has been difficult, especially since the studied diseases are rare. Not only is the problem compounded because they are based on populations with narrow ranges of exposure, but exposure misclassification can bias the results towards the null. Thus there is need for a range of exposures in the population in the study. The problems of exposure assessment need to be addressed as described above.

4. A problem with the early case-control EMF epidemiological studies was control selection bias [(5)]. In case-control studies, controls should be selected to match as closely as possible the cases under study for characteristics related to the disease excluding exposure to EMF. The participation rate should be high in both cases and controls and the approach used for selecting the controls should be well described and not be likely to introduce any bias

5. Researchers should report the basic data on which conclusions are reached, even if sophisticated statistical analyses are employed. As a minimum, the number of exposed and unexposed cases and controls in a case-control study and the number of cases observed and expected in a cohort study should be provided. Tabulations by time since exposure began and other temporal factors are also important. In a case-control study, the effects of any factors other than exposure should also be reported. When investigating cancer in a cohort study, data from all cancer sites and all causes of death should be given to reveal the possibility of reporting bias.

6. Statistical methods used to obtain absolute rates of cancer or other diseases, estimates of relative risk, confidence intervals and significance tests, and to adjust for confounding, should be clearly identified

by the researchers. Any multiple comparisons and statistical methods used should be those that are appropriate for the experiment.

Animal Studies

All known human carcinogens studied adequately in experimental animals have produced positive results in one or more animal species[(2)]. In general, if adequate data are absent from human studies, it is biologically plausible and prudent to regard studies that provide sufficient evidence of carcinogenicity in animals, as evidence of carcinogenic risk in humans[(2)]. However, the animal models need to be relevant to cancers reported in humans. The possibility that EMF may cause cancer through a species-specific mechanism which does not operate in humans should also be considered. Consistency of positive results using a variety of animal models is important.

An assessment of disease from exposure to EMF involves several considerations of qualitative importance. These include the experimental conditions under which the study was performed (exposure regimen, animal species, strain, sex, age, and duration of follow-up), the consistency of the results across species and target organs, spectrum of disease outcomes (eg for cancer, the spectrum of neoplasm response from preneoplastic lesions and benign tumours to malignant neoplasms), and the possible role of modifying factors.

Complete characterisation of EMF exposure and related environmental factors is essential for animal studies. Good laboratory practice[(8)] suggests that factors, such as exposure, animal care, pathology and statistical analyses, should be checked by an independent quality control unit and a report of their findings provided for inclusion in the final publication.

Since the probability that a disease will occur may depend on the species, sex, strain, age of the animal, and the duration of exposure, evidence of an increase in disease with level of exposure strengthens the inference of a causal association. The form of the dose-response relationship is important and may vary widely. For carcinogenesis, both DNA damage and increased cell division are important aspects.

Statistical Analysis

If human studies suggest, for example, a 25% increase in a rare cancer, the animal studies should be sensitive enough to detect this small effect. The animal model should be sufficiently well characterised so that the basic level of cancer incidence is known, and that it is low enough to detect small increases from exposure to EMF, if they occur.

When considering statistical analyses of long-term animal experiments, adequate information should be given for each treatment group. These include the numbers of animals studied and the number examined histologically, the distribution of disease types, and survival time. Types of analyses and statistical methods used should be those generally appropriate and refined for this purpose[(9)].

EVALUATION OF THE SCIENTIFIC LITERATURE

Literature for review should have been published in scientific, peer reviewed journals. Reports passing peer review should be free of most common deficiencies in methodology, analysis and conclusions. Unfortunately, the rigour of peer review varies widely among scientific journals. While peer-review adds confidence in the study results, for health risk assessment, additional review is necessary to evaluate study design, conduct and analysis of each report, and to compare them with the results of other studies. Peer-reviewed reports not published in scientific journals may be considered, but conference abstracts are of little value in health risk assessment as they generally receive no prior peer review, contain sparse information useful for a proper evaluation, and cannot be considered as the final outcome of an experiment until all results are available and properly analysed.

Criteria for Acceptance

Certain criteria should be met if individual studies reporting positive or negative effects are to be accepted into the body of established scientific literature. These criteria should be viewed as a whole; no individual criterion is either necessary or sufficient for the conclusion that there is a causal relationship between exposure and a disease.

1. Study techniques, methods and conditions should be as completely objective as possible using methodology or biological systems appropriate to end points studied. Safeguards such as double blind techniques, blind scoring or codes should be employed. Within every study there should be appropriate corresponding controls. The sensitivity of the study should be adequate to ensure a reasonable probability that an effect would be detected, if indeed any exists.

2. All data analyses should be fully and completely objective, no relevant data deleted from consideration and appropriate analytical methods used. Data from experiments within the same study should be internally consistent, within normal statistical variability. Where data are reported as ratios, the underlying data should be reported as well, or available for in-depth analysis.

3. The published description of methods should be given in sufficient detail that a critical reader would be convinced that all reasonable precautions were taken to meet requirements 1 and 2.

4. Results should demonstrate an effect of the relevant variable at a high level of statistical significance ($p>0.05$) using appropriate tests.

ASSESSMENT OF HEALTH RISK

Biological Effect Versus Health Hazard

In its constitution WHO defines health as the state of complete physical, mental and social well-being and not merely the absence of disease or infirmity. Criteria are needed to identify which EMF-induced biological effects are then to be considered a hazard to human health. Living systems respond to many stimuli as part of the process of living: such responses are examples of biological effects. The fact that a biological change is observed or suspected to occur in humans, does not by itself indicate that the environment which produces the change is hazardous. Some biological effects are inconsequential; neither hazardous or beneficial. The time course of the effect should be determined, i.e. under what conditions the effect disappears after cessation of exposure, or if exposures are additive even after a rest period, or whether effects are permanent, such as the induction of cancer.

Interactions leading to measurable biological effects which remain within the range of physiological compensation of the body and do not detract from the physical and mental well-being of humans, should not be considered as hazardous. Interactions which lead to biological effects outside the normal range of compensation of the body may be an actual or potential health hazard. If it is determined that certain EMF exposure conditions exist which have a finite probability of being unsafe for a very small population of particularly sensitive individuals, this should be addressed.

Reports of subjective effects (symptoms without concomitant signs - reactions that are difficult to measure quantitatively, e.g. headaches) are useful for identification of health consequences only if the studies are conducted in a truly scientific manner, are shown to be statistically significant and a direct causal relationship is demonstrated. Subjective effects, if substantiated, can detract from the physical and mental wellbeing of a person, and should be considered as a health hazard.

Factors in Assessing Health Risk

How can scientists evaluate the confusing and contradictory laboratory and epidemiological studies? Hill[10] developed a set of criteria that have been widely accepted when evaluating epidemiological studies. These have been elaborated further by Miller[11] and Repacholi and Stolwijk[12], and have been incorporated into the assessment of the scientific literature by WHO[13,14]. Under these criteria, strength and consistency of the association between EMF exposure and biological effects, evidence of a dose-response relationship, evidence provided by laboratory studies, and plausibility that biological systems exposed to EMF fields manifest biological effects, are all examined.

When evaluating a database for risk of cancer, or for any other health outcome from EMF epidemiological studies, the following questions need to be addressed:

1. The strength of association between exposure and risk: is there a clearly associated risk with exposure? A strong association is one with a risk ratio (RR) of 5 or more. For tobacco smoking, many of the RRs were in excess of 10. However, the EMF studies of 50/60 Hz exposures, for example, suggest a RR of about 1.5 for childhood leukaemia [5]. This is a weak association, which is more susceptible to bias and confounding than stronger associations, and alone suggests that more evidence is needed to reach any valid conclusions. Supporting evidence of cancer in laboratory animals exposed to EMF fields would be important to increase confidence that the epidemiological studies could be indicating a real risk.

2. How consistent are the studies of association between exposure to EMF fields and the risk of cancer? Do most studies show the same risk for the same disease? Using the example of smoking, essentially all epidemiological studies of smoking demonstrated an increased risk for lung cancer. Studies may show statistically significant associations between some types of cancers and some types of exposures, but others do not. Alternatively, studies reporting an association between cancer may be inconsistent with each other in their types or subtypes. The ability of the study design to identify true risk without bias and confounding should be weighed.

3. Is there a dose-response relationship between exposure to EMF fields and the risk of cancer? Again, the more a person smokes, the higher the risk of lung cancer. Do the EMF field exposure studies demonstrate a dose-response relationship between measured, calculated, or estimated EMF fields and cancer rates?

4. Is there laboratory evidence for an association between exposure to EMF and the risk of cancer? When warnings that smoking caused lung cancer first appeared, the epidemiological evidence was very strong but the laboratory evidence was ambiguous. It was known that cigarette smoke and tobacco contained carcinogens, but no study had demonstrated cancer from smoking in laboratory animals. This problem has now been overcome and laboratory evidence linking smoking to cancer is stronger. Thus, the evidence is considered much stronger if effects can be demonstrated in animals rather than cells or tissues alone, since whole animals are able, through various mechanisms, to amplify, minimise or negate the effects of exposure to physical agents. The weight assigned to studies of whole animals is greater than the weight assigned to studies of isolated tissues and cells because of the absence of systemic regulatory controls and mechanisms in cells and tissues.

5. Are there plausible biological mechanisms for a link between EMF field exposure and the risk of cancer? When it is understood how an agent causes disease, it is easier to interpret ambiguous epidemiological evidence and to design better and more powerful epidemiological studies. For smoking, while the direct laboratory evidence connecting smoking with cancer was initially weak, the association was highly plausible because there were known cancer causing agents in tobacco smoke. The biological significance of responses observed *in vitro* should not be assumed unless it has been demonstrated that similar responses do occur in vivo and are relevant to human health effects.

Evaluation of Carcinogenicity

Assessment of health effects such as cancer will receive special attention within the International EMF Project as there are many reports that exposure to EMF fields may be associated with increased cancer risk. Evaluations of the strength of evidence for carcinogenicity arising from human and animal data will be based on the criteria developed by the IARC[2]. However, it has been noted that the Environmental Protection Agency[15] have released draft guidelines for comment on the procedures for assessing carcinogenesis. EPA suggests placing more weight on mechanisms of action. The procedures to be used in the International EMF Project for evaluating cancer risk from EMF exposure have been elaborated by Cardis and Rice[3].

Within the International EMF Project, final assessments of health risk will be made by formally constituted WHO Working Groups comprising scientists from all appropriate disciplines, with representation by gender and from various geographical regions. Working Group members are appointed by the Executive Director of WHO's Programme on Environment and Health.

IARC[2] assigns categories related to degrees of evidence for carcinogenicity in humans and experimental animals. These categories refer only to the strength of evidence that exposure is carcinogenic and not to the extent of its carcinogenic activity (potency) nor to the mechanisms involved. A classification may change as new information becomes available.

Carcinogenicity in Humans

The applicability of an evaluation of carcinogenicity of exposure in given situations, occupations or industries on the basis of evidence from epidemiological studies depends on the variability over time and place of exposure. The Working Group will identify the specific exposure or activity which is considered most likely to be responsible for any excess health risk. The evidence relevant to carcinogenicity from studies in humans is classified into one of the categories: given below. In some instances, these categories may be used to classify the degree of evidence related to carcinogenicity in specific organs or tissues.

Sufficient evidence of carcinogenicity. The Working Group considers that a causal relationship has been established between exposure and human cancer. That is, a positive relationship has been observed between the exposure and cancer in studies in which chance, bias and confounding could be ruled out with reasonable confidence.

Limited evidence of carcinogenicity. A positive association has been observed between exposure and cancer for which a causal interpretation is considered by the Working Group to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence.

Inadequate evidence of carcinogenicity. The available studies are of insufficient quality, consistency or statistical power to permit a conclusion regarding the presence or absence of a causal association, or no data on cancer in humans are available.

Evidence suggesting lack of carcinogenicity. There are several adequate studies covering the full range of levels of exposure that human beings are known to encounter, which are mutually consistent in not showing a positive association between exposure to EMF and any studied cancer at any observed level of exposure. A conclusion of "evidence suggesting lack of carcinogenicity" is inevitably limited to the cancer sites, conditions and levels of exposure and length of observation covered by the available studies. In addition, the possibility of a very small risk at the levels of exposure studied can never be excluded.

Carcinogenicity in Experimental Animals

Evidence relevant to carcinogenicity in experimental animals is classified into one of the following categories:

Sufficient evidence of carcinogenicity. The Working Group considers that a causal relationship has been established between exposure and an increased incidence of malignant neoplasms or of an appropriate combination of benign and malignant neoplasms in (a) two or more species of animals or (b) in two or more independent studies of one species carried out at different times or in different laboratories or under different protocols. Exceptionally, a single study of one species might be considered to provide sufficient evidence of carcinogenicity when malignant neoplasms occur to an unusual degree with regard to incidence, site, type of tumour or age at onset.

Limited evidence of carcinogenicity. The data suggest a carcinogenic effect but are limited for making a definitive evaluation because, e.g. (a) the evidence of carcinogenicity is restricted to a single experiment; or (b) there are unresolved questions regarding the adequacy of the design, conduct or interpretation of the study; or (c) exposure increases the incidence only of benign neoplasms or lesions of uncertain neoplastic potential, or of certain neoplasms which may occur spontaneously in high incidence in certain strains.

Inadequate evidence of carcinogenicity. The studies cannot be interpreted as showing either the presence or absence of a carcinogenic effect because of major qualitative or quantitative limitations, or no data on cancer in experimental animals are available.

Evidence suggesting lack of carcinogenicity. Adequate studies involving at least two species are available which show that, within the limits of the tests used, exposure is not carcinogenic. A conclusion of evidence suggesting lack of carcinogenicity is inevitably limited to the species, tumour sites and levels of exposure studied.

Other Data Relevant to the Evaluation of Carcinogenicity

Other evidence judged to be relevant to an evaluation of carcinogenicity and of sufficient importance to affect the overall evaluation is also considered. This may include data on preneoplastic lesions, tumour pathology, genetic and related effects, structure-activity relationships, metabolism, physicochemical parameters and analogous biological agents.

Data relevant to mechanisms of the carcinogenic action are also evaluated. The strength of evidence that any carcinogenic effect observed is due to a particular mechanism is assessed, using terms such as weak, moderate or strong. The Working Group then assesses if the particular mechanism is likely to be operative in humans. The strongest indications that a particular mechanism operates in humans come from data on human or biological specimens obtained from exposed humans. Data may be considered to be especially relevant if they show that exposure in humans has caused changes that are on the causal pathway to carcinogenesis.

Overall Evaluation

Finally, the body of evidence is considered as a whole, in order to reach an overall evaluation of the carcinogenicity to humans. A common approach for determining this is by weight of evidence. There is no way to prove something does not cause cancer since no foolproof test exists for carcinogens or hazard identification. Thus it is necessary to estimate how much of a given set of evidence (established scientific database) changes the probability that exposure is carcinogenic.

The carcinogenicity of exposure is described according to the wording of one of the following categories. The categorization of exposure is a matter of scientific judgement, reflecting the strength of the evidence derived from studies in humans, animals and from other relevant data.

Group 1 - Exposure is carcinogenic to humans.

This category is used when there is sufficient evidence of carcinogenicity in humans. Exceptionally, exposure may be placed in this category when evidence in humans is less than sufficient but there is sufficient evidence of carcinogenicity in experimental animals and strong evidence in humans that exposures act through a relevant mechanism of carcinogenicity.

Group 2

This category includes exposure for which, at one extreme, the degree of evidence of carcinogenicity in humans is almost sufficient, as well as those for which, at the other extreme, there are no human data but for which there is evidence of carcinogenicity in experimental animals. Exposure is assigned to either group 2A (probably carcinogenic to humans) or group 2B (possible carcinogenic to humans) on the basis of epidemiological and experimented evidence of carcinogenicity and other relevant data.

Group 2A - Exposure is probably carcinogenic to humans.

This category is used when there is limited evidence of carcinogenicity in humans and sufficient evidence of carcinogenicity in experimental animals. In some cases, exposure may be classified in this category when there is inadequate evidence of carcinogenicity in humans and sufficient evidence of carcinogenicity in experimental animals and strong evidence that the carcinogenesis is mediated by a mechanism that also operates in humans. Exceptionally, exposure may be classified in this category solely on the basis of limited evidence of carcinogenicity in humans.

Group 2B - Exposure is possibly carcinogenic to humans.

This category is used when there is limited evidence of carcinogenicity in humans and less than sufficient evidence of carcinogenicity in experimental animals. It may also be used when there is inadequate evidence of carcinogenicity in humans but there is sufficient evidence of carcinogenicity in experimental animals. In some instances, if there is inadequate evidence of carcinogenicity in humans but limited evidence of carcinogenicity in experimental animals, together with supporting evidence from other relevant data, exposure may be placed in this group.

Group 3 - Exposure is not classifiable as to its carcinogenicity to humans.

This category is used most commonly when the evidence of carcinogenicity is inadequate in humans and inadequate or limited in experimental animals. Exceptionally, if there is inadequate evidence of carcinogenicity in humans but sufficient in experimental animals, exposure may be placed in this category when there is strong evidence that the mechanism of carcinogenicity in animals does not operate in humans.

Group 4 - Exposure is probably not carcinogenic to humans

This category is used when there is evidence suggesting lack of carcinogenicity in humans and in experimental animals. In some instances, if there is inadequate evidence of carcinogenicity in humans but evidence suggesting lack of carcinogenicity in experimental animals, and this is consistently and strongly supported by a broad range of other relevant data, exposure may be classified in this group.

CONCLUDING REMARKS

This paper indicates the type of research (ie characteristics of a scientific database) needed to assess health risk, the basis by which literature reviews are conducted to reach scientifically valid conclusions, and the criteria to assess health risk from exposure to EMF fields within in the International EMF Project. Details on progress of the International EMF Project can be found on its home page at: http://www.who.ch/programmes/peh/emf/emf_home.htm.

REFERENCES

1. Repacholi, *M.H. The international electromagnetic fields (EMF) project.* In: R Matthes (ed) Non-ionizing radiation. Proceedings of the third international non-ionizing radiation workshop, Baden.(1996). Available from: R Matthes, Scientific Secretary, ICNIRP, C/- Bundesamt für Strahlenschutz, Institut für Strahlenhygiene, Ingolstädter Landstraße 1, D-85764 Oberschleißheim, Germany.

2. International Agency for Research on Cancer. *IARC Monographs on the evaluation of carcinogenic risks in humans: Preamble.* Lyon, International Agency for Research on Cancer (1995).

3. Cardis, E and Rice, J.M. *Criteria for health risk assessment within the International EMF Project.* In: Bernhardt, J.H, Matthes, R, Repacholi, M.H. (eds) Non-thermal effects of RF electromagnetic fields. Proc. International seminar Biological effects of non-thermal pulse and amplitude modulated RF electromagnetic fields and related health hazards. Munich-Neuherberg, 20-22 November, 1996. pp 37-46 (1997). Available from: Scientific Secretary, International Commission on Non-Ionizing Radiation Protection, C/- Bundesamt für Strahlenschutz, Institut für Strahlenhygiene, Ingolstädter Landstraße 1, D-85764 Oberschleißheim, Germany.

4. Wertheimer, N and Leeper, E. *Electrical wiring configurations and childhood cancer.* Am. J. Epidemiol. 109(3) 272-284 (1979).

5. NRC *Possible health effects of exposure to residential electric and magnetic fields.* National Research Council, Washington: National Academy Press (1996).

6. Feychting,M and Ahlbom,A. *Magnetic fields and cancer in children residing near Swedish high voltage power lines.* Am. J. Epidemiol. 138 467-481 (1993).

7. Repacholi, M.H. (ed.) *Low-level exposure to radiofrequency fields: Health effects and research needs.* Bioelectromagnetics (1997)(in press).

8. FDA *Good laboratory practice for nonclinical laboratory studies.* Food and Drug Administration, US Department of Health and Human Services. Federal Register 21 CFR Ch. 1 (4-1-93 Edition) Part 58, pp 245-258 (1993).

9. Gart, J.J, Krewski, D, Lee, P.N, Tarone, R.E and Wahrendorf, J. *Statistical methods in cancer research,* Vol. 3, The design and analysis of long-term animal experiments (IARC Scientific publications No. 79), Lyon, IARC (1986).

10. Hill, A.B. *The environment and disease: Association or causation?* Proceedings of the Royal Society of Medicine 58 295-300 (1965).

11. Miller, R.W. *Aetiology and epidemiology* In: Cancer in children - clinical management. Voute, P.A , Barrett, A, Bloom, H.J.G, Lemerle, J, Neidhardt, M.K. (eds) Berlin: Springer-Verlag pp 3-8 (1986).

12. Repacholi, M.H. and Stolwijk, J.A.J. Criteria for evaluating scientific literature and developing exposure limits. Radiation Protection in Australia, 9(3) 79-84 (1991).

13. WHO *Magnetic Fields*, Environmental Health Criteria N 69, Geneva, World Health Organization, (1987).

14. WHO *Electromagnetic Fields (300 Hz - 300 GHz)*, Environmental Health Criteria N 137, Geneva, World Health Organization (1993).

15. Environmental Protection Agency. *Proposed guidelines for carcinogen risk assessment; Notice.* Federal Register 61 (79) 17961-18011(1996).

REFERENCES

1. Portier CJ, Wolfe MS, eds. Assessment of Health Effects from Exposure to Power-Line Frequency Electric and Magnetic Fields. Research Triangle Park NC: National Institute of Environmental Health Sciences, National Institutes of Health, 1998.
2. National Academy of Sciences, National Research Council. Possible Health Effects of Exposure to Residential Electric and Magnetic Fields. Washington DC: National Academy Press, 1996.
3. Adair R Biological responses to weak 60 Hz fields must vary as the square of the field strength. *Proceedings of the National Academy of Sciences* 91: 9422-5,1994
4. Wilson R, Risks assessment of EMF on health, *IEEE Engineering in Medicine and Biology*, July/August 1996, pp77-86.
5. Greenland S. Probability, logic and probabilistic induction. *Epidemiology* 9:322-2, 1998
6. Ockham's razor. In: Encyclopedia Britannica. Chicago IL, 1998.
7. Huff J. Chemicals and cancer in humans: First evidence in experimental animals. *Environmental Health Perspectives* 100:201-10, 1993.
8. Gold LS, Slone TH, Stern BR. Rodent Carcinogens: Setting priorities. *Science* 258:261-5, 1992
9 Fung VA et al. The carcinogenesis bioassay in perspective: application in identifying human cancer hazards. *Environmental Health Perspectives* 103:680-3, 1995.
10. Hill B. The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine* 58:295-300, 1965
11. Rothman K, Greenland S, eds. Modern Epidemiology 2nd Edition. Philadelphia PA: Lippincott-Raven, 1998.
12. Lindbohm M-L, Hietanen M, Pentti K et al. Magnetic fields of video display terminals and spontaneous abortion. *American Journal of Epidemiology* 136:1041-51, 1992.
13. Linet M, Hatch E, Kleinerman R. Residential exposure to magnetic fields and acute lymphoblastic leukemia in children. *New England Journal of Medicine* 337:1-7, 1997.
14. Valberg PA. Designing EMF experiments: What is required to characterize "exposure"? *Bioelectromagnetics* 16:396-401 (discussion 402-6), 1995.
15. Zaffanella L. Survey of Residential Magnetic Field Sources. Volume One and Volume Two TR-102759. Palo Alto, CA: EPRI, 1993.
16. McBride ML, Gallagher RP, Thériault G, et al. Power-frequency electric and magnetic fields and risk of childhood leukemia in Canada. *American Journal of Epidemiology* 149:831-42, 1999.
17. London SJ, Preston-Martin S. EMF exposure and breast cancer in minority women in Los Angeles. In preparation, 1999.
18. Lee GM, et al. EMF exposure and spontaneous abortion in the San Francisco Bay area. In preparation, 1999.
19. International Agency for Research on Cancer (IARC) monograph programme on the evaluatio of the carcinogenic risk of chemicals to humans. pp. 15-35, 1978.
20. Zeller RA, Carmines EG, eds. Measurements in the Social Sciences: The Link between Theory and Data. New York NY: Cambridge University Press, 1980.
21. Jensen FV. An Introduction to Bayesian Networks. Springer, NY, 1996
22. Hutchison T, Lane D. Assessing method for causality assessment of suspected adverse drug reactions. *Journal of Clinical Epidemiology* 42:5-16, 1980.